

Predicting Students' Academic Performance using Artificial Neural Networks

Solomon Tok Dung

Department of Computer Science,
National Open University of Nigeria
Jos study centre

Godwin Thomas

Department of Computer Science,
University of Jos

Yinka Oyerinde

Department of Computer Science,
University of Jos

ABSTRACT

Currently, within our contemporary society and the world at large, it is next to impossible not to find technology in almost all sectors of human life, the educational sector inclusive. As part of policies in any educational institution, improving students' academic performance is paramount amongst other expectations. In order to accomplish this, most secondary schools rely on the manual method of assumption and MOCK exams to evaluate students' progress with the view of improving their outcome. This approach has been used over the years and it is evident that it has proven to be ineffective and inefficient. The proof can be seen in the continuous dwindling on the WAEC results of students in both private and public secondary schools in Nigeria. Now, despite the availability of Information and Communication Technology (ICT) and also students' data, most secondary schools do not use the potentials embedded in the data available to look for a solution to this problem. Given existing approach(es) limitations to improving students' academic performance, the need for a better approach arises. Thus, this research proposes an approach that leverages on AI technology to realize better results. The research focuses on the use of ICT and Artificial Neural Networks to aid in predicting students' academic performance. By employing both qualitative and quantitative approaches the requirements for such a system were identified. WEKA was used as the simulation environment to test the proposed algorithm premised on serial factors identified. The test result premised on a prototype, shows that the model can be used to predict students' academic performance in secondary schools.

General Terms

Predictive analytics, predictive modeling, Artificial Neural Networks

Keywords

WEKA, academic performance data, contextual factors, sigmoid function, mixed method research design, abstraction of prototype.

1. INTRODUCTION

Neural networks (NN) are transforming industry and daily life and advancing artificial intelligence (AI). NN-enabled technologies (including the smartphones and computers that we use on a daily basis) are now trained to learn, detect patterns, and make predictions in a humanoid manner as well as solve issues in every business sector by simulating the way interconnected brain cells function (Ramos, 2018). Neural networks are ideally suited for today's big data-based applications because of their human-like characteristics and capacity to execute jobs in an endless number of permutations and combinations. When perfect models are not available, neural networks can still use controlled processes because they have the special ability (known as fuzzy logic) to make sense

of confusing, contradictory, or incomplete input. This type of application is visible in the educational sector.

Prediction of students' academic performance is one of the educational problems solved by data mining. Many researchers have developed models for predicting student performance at various levels using various Data Mining (DM), Machine Learning (ML), and Statistical methods based on various data (Cortez and Silva, 2008; Hamsa et al., 2016; Oyerinde and Chia, 2017; Mubarak, 2019). In the majority of these studies, neural networks outperformed other methods, and none of the methods could discover potential data patterns as well as neural networks. Because of the positive results of using neural networks in prediction and classification problems, it is appropriate for this study.

The problem explored in this study is that secondary schools in Nigeria still rely on the traditional methods of assumption in predicting students' academic. This makes adequate intervention to improve their outcomes difficult and ineffective as students' performance continues to dwindle. The solution proposed in this research is to develop a predictive model for the schools which will be able to predict the students' final results, thereby eliminating the reliance of teachers on assumptions and focus on students grades only as other factors have been established to also affect the performance of students. It will also be able to identify the key areas which need to be improved on for at a glance for adequate intervention on the part of the school with the sole purpose of enhancing their outcomes.

This study utilizes Artificial Neural Networks to predict the performance of secondary school students in their final year exams using the average of their past grades in various subjects in SS1, SS2 and SS3, in addition to other variables known to affect students' performance prior to them sitting for the examinations. It has been recognized that failure has continuously permeated the performance of students within the progression. Therefore, by utilizing Neural Networks, this research work aims to address this issue to help identify the weak spots from the nascent stage. This has become pertinent as we look to discover patterns hidden in the decline and dwindling academic performance of students in both private and public schools with regards to their final year examinations: in this case Senior Secondary Certificate Examinations (SSCE), a prerequisite to obtaining admission into colleges, polytechnics and universities for tertiary education. This is similar to the SAT or GED in western countries. The next section gives more light on the dwindling nature of the students' performance over the range of some years.

2. RELATED CONCEPTS

2.1 Predictive Analytics and Modeling

In general, predictive analytics refers to analytics which carry out forecasting or extrapolation in order to build accurate models that predict future occurrences or results from past occurrences (Oyerinde, 2019). However, the term is increasingly being used to refer to related analytical disciplines such as descriptive modeling, decision modeling, and optimization. Predictive analytics is a branch of statistics concerned with extracting information from data and using it to forecast trends and behavioral patterns. Predictive web analytics enhancement calculates statistical probabilities of future events online. Data modeling, machine learning, AI, deep learning algorithms, and data mining are examples of predictive analytics statistical techniques. The unknown event of interest is frequently in the future, but predictive analytics can be applied to any type of unknown, whether in the past, present, or future. Identifying suspects after a crime has been committed, for example, or credit card fraud as it occurs. Predictive analytics is based on capturing relationships between explanatory variables and predicted variables from previous occurrences and using them to predict an unknown outcome. It is important to note, however, that the level of data analysis and the quality of hypotheses will have a significant impact on the accuracy and usability of the results.

2.2 How Neural Networks are Arranged

The Artificial Neural Network is modeled similarly to the structure of the Natural Neuron. As such, it has similarities with the human brain. Figure 1 below shows a natural neuron comprising of a nucleus, dendrites, and axon extending itself into various branches to form connections (synapses) with other neurons via dendrites.

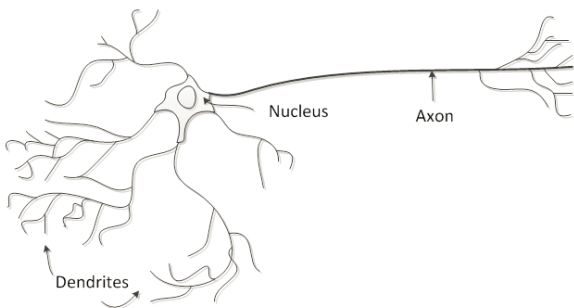


Figure 1: Natural Neuron

Similarly, the artificial neuron has the same structure. It consists of a nucleus (processing unit), dendrites (inputs), and an axon (output).

2.3 Artificial Neuron

Natural neurons can be said to be signal processors, since they receive small signals from the dendrites that trigger a signal in the axon depending on their strength or magnitude. Suffice it to say that neuron has a signal collector in the inputs and an activation unit in the output that triggers a signal which will be forwarded to other neurons. Figure 2 below gives a pictorial description of the Artificial Neuron

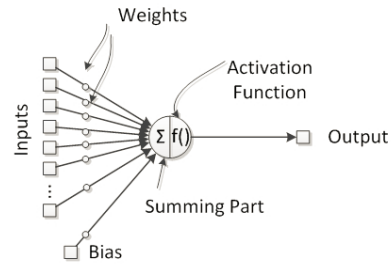


Figure 2: Artificial Neuron structure

2.3.1 Activation Function

The activation function provides the Output of the Neuron. It adds nonlinearity to neural network processing, which is essential to emulate the nonlinear attributes of a natural neuron. The output of an activation function is usually bounded between two values, a characteristic of a nonlinear function. However, it can be a linear function. The four most used activation functions are as follows:

1. Sigmoid
2. Hyperbolic tangent
3. Hard limiting threshold
4. Purely linear

The activation function utilized in this model is the sigmoid function. Below is a brief explanation on its mode of operation.

2.3.1.1 Sigmoid function

The logistic function is another name for the sigmoid activation function. The function outputs numbers between 0 and 1 and accepts any real value as input. The output value will be nearer to 1.0 the larger the input (more positive), and nearer to 0.0 the smaller the input (more negative).

This is how the sigmoid activation function is calculated:

$$\bullet 1.0 / (1.0 + e^{-x})$$

where the natural logarithm's base, e , is a mathematical constant.

The outputs for a variety of values and plotting an input-to-output graph we can see the sigmoid activation function's well-known S-shape.

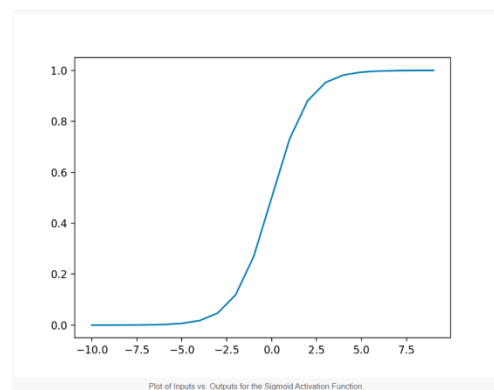


Figure 3: Graph of the sigmoid function

2.3.2 Weights

Weights can be described as connections between neurons with the ability to modify neuron signals. As such, they have an influence on the output of the neuron, as a neuron's activation depends on the inputs and on the weights. Provided that the inputs come from other neurons or from the external world, the

weights are considered to be a neural network's established connections between its neurons. Thus, since the weights are internal to the neural network and influence its outputs, we can consider them as neural network knowledge, provided that changing the weights will change the neural network's capabilities and therefore actions.

2.3.3 Bias

The **bias**, in an artificial neuron, is an independent component which adds an extra signal to the activation function. They have an associated weight which helps in the neural network knowledge representation as a more purely nonlinear system.

2.3.4 Layers

Natural neurons are typically organized in layers. Each layer provides a specific task. The Layers can be basically divided into three classes:

Input layer
Hidden layer
Output layer

The input layer receives direct stimuli from the outside world, and the output layers fire actions that will have a direct influence on the outside world. Between these layers, there are a number of hidden layers, in the sense that they do not interact directly with the outside world.

2.4 The Conceptual ANN Frame

The main aim of this research is to develop and construct a model for Predicting Students Academic Performance using Artificial Neural Networks. This section presents the conceptual frame from which the proposed model is predicated. Figure 4 presents the conceptual frame for the proposed model.

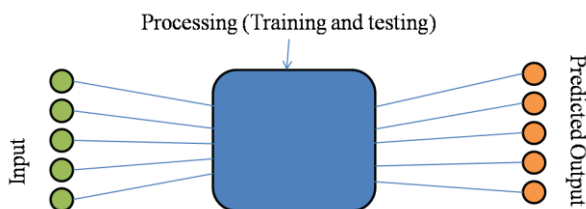


Figure 4: The Conceptual ANN frame

There are basically three stages involved. The description of the stages utilized in this project which are input design, process design and output design are presented in sections 2.4.1-2.4.3.

2.4.1 Input Design

The type of data to be taken as input includes:

1. Students' academic information.
2. Subjects offered by the students.
3. Students' exam scores in each subject they are taking per term for three years.

Numerical values of the factors influencing the performance of the students in each subject.

2.4.2 Process Design

For this research which utilizes Neural Networks, the under-listed are considered in the process design:

1. Preprocessing of data

2. Selecting the function
3. Number of layers
4. Learning rate
5. Momentum
6. Number of iterations

2.4.3 Output Design

The expected outputs from the system include:

1. Predicted scores of the students according to each subject offered.
2. Mean Square Error of the Neural network
3. Correlation coefficient of the attributes.
4. Accuracy of the prediction.

3. DESIGN METHODOLOGY

In order to get the requirements for this study, we will need more than just data from students' academic performance data as their performance can also rely on other factors as well which are contextual factors. Thus, to identify the relevant factors that will play a role in determining a student's potential performance this study employs "Mixed Method Research Design to determine relevant factors. Where after a prototype is instantiated to show a proof of concept. The mixed method approach employed to identify the proposed ANN model input factors is further elaborated in the following subsections.

3.1 Mixed Method Research Design

To better understand a research problem, a mixed methods research design is employed. A Mixed-method research design is a research approach that combines qualitative and quantitative research methods in a single study to provide a comprehensive understanding of a research problem (Creswell, 2012). Mixed-method research is becoming increasingly popular because it allows researchers to draw on the strengths of both qualitative and quantitative methods, while also compensating for their weaknesses.

This study employs a mixed methods research design, combining both quantitative and qualitative data to develop an artificial intelligence program to predict students' future academic performance. The quantitative component of the study involved collecting academic performance data for students, including their grades and test scores, from school records. The qualitative component of the study involved conducting semi-structured interviews with teachers and students to collect data on contextual factors that may influence academic performance, such as family background, motivation, and study habits.

The collected data was then used to develop an artificial intelligence program, which integrated both the quantitative and qualitative data to predict students' future academic performance. The program was trained using machine learning algorithms, with the academic performance data serving as the input variables, and the contextual factors identified through interviews serving as additional input variables. The output of the program was a predicted academic performance score for each student.

The program's performance was evaluated using a hold-out sample of students, and the accuracy of its predictions was compared to traditional methods of predicting academic performance, such as WAEC result. The results showed that the artificial intelligence program outperformed these traditional methods, indicating that the integration of both quantitative and qualitative data can improve the accuracy of predictions of

academic performance. The quantitative and qualitative research contexts are elaborated in the following subsections.

3.1.1 Quantitative Research

A style of educational research in which the subject matter is chosen in advance, a focused set of questions are posed, a large number of participants provide quantifiable data, which is then analyzed statistically, and the investigation is carried out impartially and objectively. The variables derived in this context can also be referred to as academic performance data. An example is the scores or grades of the students for each subject. The table1 below illustrates further on this.

Table 1: Quantitative or academic performance data

Subject	Grades		
	Year 1	Year 2	Year 3
English	55	74	70
Maths	75	72	54
Physics	75	70	74
Chemistry	69	75	75
Biology	75	78	74
Geography	75	70	80
Civic	75	80	82
Computer	75	75	76
Agric	92	75	80

3.1.2 Qualitative research (Contextual factors)

A kind of educational research where the researcher relies on the opinions of participants, asks open-ended, general questions, gathers information primarily from participants' words (or texts), describes and examines these words for themes, and conducts the investigation in a biased, subjective manner. The variables identified from this perspective are referred to as contextual variables(attributes). Table 2 below is an illustration of the qualitative or perception experts variables with examples of assigned score values out of a maximum value of 10 in positive situation.

Table 2: Qualitative or Contextual factors variables

Subject	Contextual factors			
	Student factor	School factor	Societal factor	Parental factor
Maths	9	8	3	5
Physics	8	3	4	8
Chemistry	5	5	4	2
Biology	6	7	5	2
Geography	8	7	8	8
Civic	9	8	8	9
English	9	8	8	9

3.2 The ANN model prototype instance

This section aims to provide an instance of the ANN model as a proof of concept, showing its potential in terms of predicting students' academic performance. For the implementation of the model in this research, the Waikato Environment for Knowledge Analysis (WEKA) software was used.

The reason for the choice of this model is due to the fact that it meets most of the requirements of the proposed system. The model has various data mining algorithms and its application is found in many different areas, particularly educational and research purposes.

The procedure to be followed with this prototype is as follows: once the students data or variables have been gathered, it has

been collated and arranged in a file as input. This input is split into two sets: Training set and Test set. The Training set comprises data of students who have already sat for the exam and the results known. The Training dataset is then fed into the model which processes the data based on predetermined conditions set in the training phase of the model to achieve optimal results. The Test set is then used as fresh data to predict the outcome of the students yet to sit for the exams.

3.3 Requirements Specifications

The requirements of any system are usually divided into two, the functional requirements and the non-functional requirements.

Functional requirements describe what the software system should do. In addition to that, the literature review, interviews and questionnaires further established additional requirements needed to effectively design the model as outlined in the next section. The non-functional requirements on the other hand place constraints on how the system will do so.

3.3.1 Functional Requirements

1. The input field should contain data comprising previous scores of students over the years in various subjects prior to the examinations (i.e average for year 1, year 2 and year 3), their final result after the examination (WAEC result in this case for the train dataset) and values for the factors affecting students performance for each student which will be scaled for the model to utilize.
2. The data should be split into two sets; train dataset and test dataset.
3. The train data set consists of the following;
 - i. Subject
 - ii. Grades (year 1,2,3)
 - iii. Factors/ contextual data (Student, School, Teacher, Parent e.t.c)
 - iv. Final result (Outcome)
4. The test dataset consists of the same as above except the final result which will be predicted.
5. The model should be able to validate the compatibility of the two data sets to prevent runtime error.
6. The system must be able to automatically predict the outcome of a fresh data set.

3.3.2 Non-Functional Requirements

1. The system must have an easy to understand and use user interface.
2. The system must allow easy navigation.
3. The system must display information in an orderly and understandable manner.
4. The system must be able to handle recovery of data in case of loss.
5. The system must allow room for new features to be added.
6. The system must be maintainable.

3.4 Database Design

The data collected was collated in Microsoft excel and saved as a Comma Separated Variable (.csv) format which is a format acceptable by WEKA. It can also be converted into Attribute Related File Format (.arff). Figure 5 below is the entity relationship diagram for the prototype following the requirements stated in the input design stage.

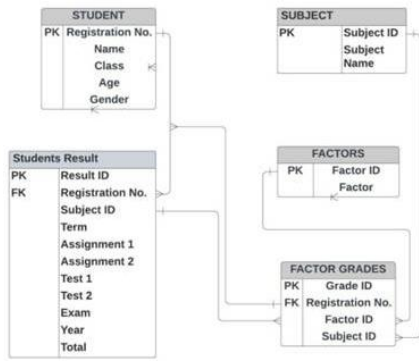


Fig 5 Entity Relationship Diagram for the prototype

3.5 Use Case Model

The figure6 shows the use case model showing how the user interacts with the system.

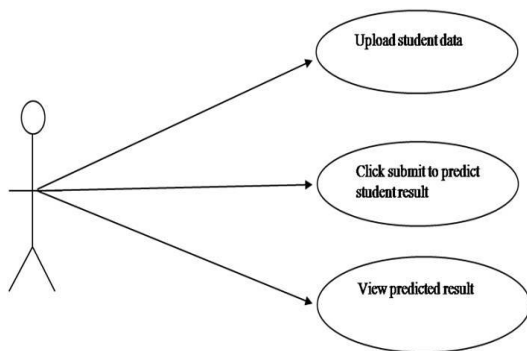


Figure 6: Use case Model

3.6 Flowchart

This is a diagram that represents an algorithm, workflow or process of the system. The flow of data is, the user uploads students data already arranged in a format acceptable by the

Neural network, clicks the predict button and then the system predicts the performance using the trained model. Figure 7 below shows the flow for the system.



Figure 7 Flowchart of the model

3.7 Model function and processes

The major function for the model can be derived from the structure of the artificial neuron as depicted in figure 2.2 in section 2. The predicted outcome can be defined by the formula

$$P = \sum f_n, \text{ where}$$

P = Predicted outcome

Σ = summation

f_n = activation function = (inputs * weights + bias)

input = (academic performance data variables and Contextual factors variables). Figure 8 illustrates the proposed model for the prototype.

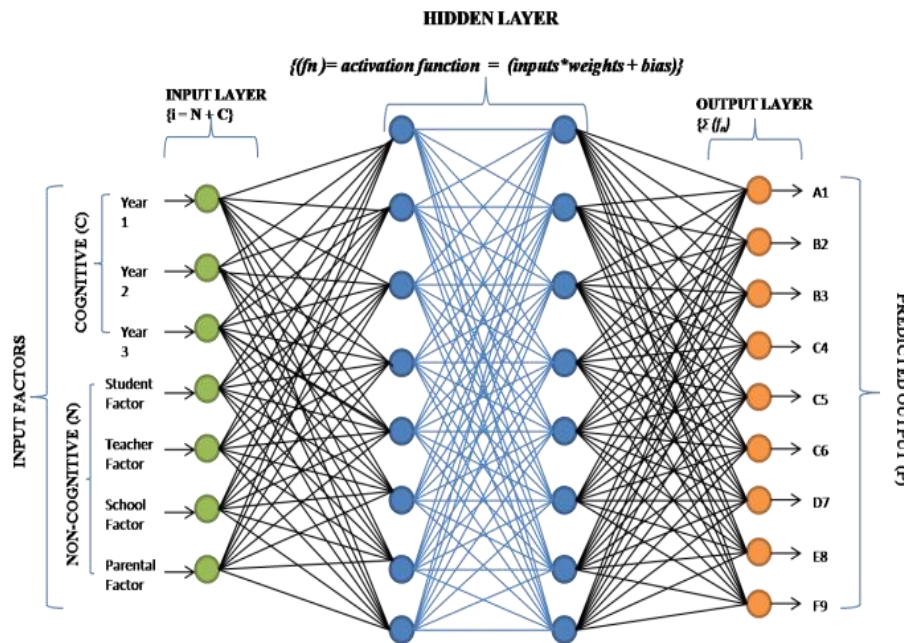


Figure 8: Abstraction of the model prototype

After students data is gathered (scores or grades as the academic performance data variables "C" then factors influencing the students performance as contextual factors

variables "N"), this data (train dataset) is fed into the model as input "i". The hidden layer handles the processing of the data by applying an activation function which is a product of the

inputs "i" and weights "w" in addition to a bias which continuously adjusts the weights till the desired output "P" is achieved. This can be denoted by the formula below

$$f_n = (i_1w_1 + i_2w_2 + \dots + i_nw_n + bias)$$

The model is then saved and used on the test dataset. Now when the inputs are high i.e both academic performance data (70-100) and contextual factors variables(7-10), the output is expected to be high(within the range of A1 - B3). When one of the variables is low and the other is high, the output becomes moderate or low (within the range of C4-D7). When both inputs are low, the output becomes low. Table 3 gives an illustration of the scenarios explained.

Table 3: Range of input with expected output

SUBJECT	ACADEMIC PERFORMANCE DATA			CONTEXTUAL FACTORS				OUTPUT
	YEAR 1	YEAR 2	YEAR 3	STUDENT FACTOR	SCHOOL FACTOR	SOCIETAL FACTOR	PARENTAL FACTOR	
ENGLISH	55	74	70	9	8	3	5	B2
MATHS	75	72	54	8	3	4	8	B3
PHYSICS	75	70	74	5	5	4	2	A1
CHEMISTRY	69	75	75	6	7	5	2	B3
ENGLISH	39	39	39	4	3	5	4	E8
MATHS	35	30	34	3	4	3	3	F9
PHYSICS	49	59	39	4	3	4	4	F9
BIOLOGY	39	35	36	3	3	5	4	F9
COMPUTER	49	40	54	5	4	3	5	E8
MATHS	35	40	49	8	5	4	10	D7

4. SYSTEM IMPLEMENTATION

The main software required to set up the system is Waikato Environment for Knowledge Analysis (WEKA) software. WEKA supports several data mining tasks, more specifically, data preprocessing, clustering, regression, visualization and feature selection. For the Artificial Neural Network in WEKA, the Multilayer perceptron is the function utilized.

For this research, the database generated was derived by employing the Mixed method research design mentioned in section 3. The Quantitative or academic performance data was gotten from students record. The Qualitative or contextual data was gotten from an evaluation of the factors perceived to affect student academic factors through the administration of questionnaire surveys both sets of data were gotten from some chosen schools both public and private. This data is characterized by instances and attributes as seen in table 3 in section 3. So, for example we have the record of 10 students offering 7 subjects each, the total number of instances will then be 70 instances with 7 attributes.

The operation of the prototype is carried out in two phases: training and testing. Recall that in our literature review found in section 2, the efficiency of any Neural Network depends on training the Neural network first before it is deployed on a test set. Hence, the need for operating it in these two phases. The datasets gotten from the questionnaires and data sheets administered to the schools was divided into the training set and testing set. The data collected was split into two having the same number of instances: training dataset and test dataset. The training set comprises the data obtained from students who

have already graduated while the test set is gotten from those yet to graduate. Figure 9 shows the Neural Network designed by the program with its corresponding nodes, layers and parameters set for the Neural Network.

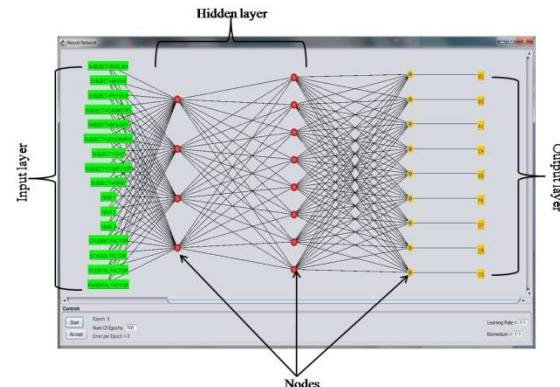


Figure 9: Neural Network GUI

In Figure 9, the **Input layer** consists of the input factors taken by the Neural Network comprising of academic performance data variables as "Grades of the subjects"(Year 1, Year 2, Year 3) and Contextual factors variables (Student factor, School factor, e.t.c) as discussed in section 3. The hidden layer is where the processing is carried out before an output is displayed on the Output layer using the sigmoid function as an activation function. The Output, as depicted in Figure 9 above is the result (i.e A1, B2, B3, C4 e.t.c).

Figure 10 depicts the result of the prediction along-side other details to enhance the evaluation of the systems performance.

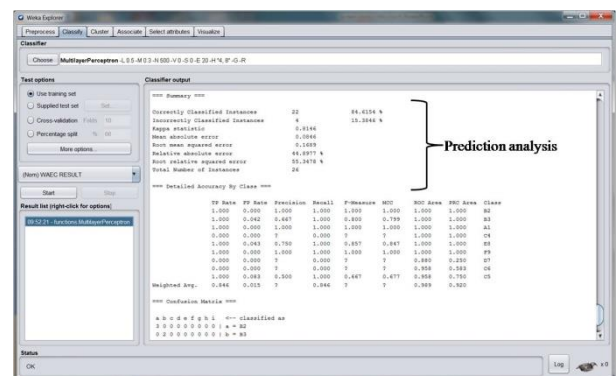


Figure 10: Prediction page

5. CONCLUSION AND FUTURE WORK

5.1 Conclusion

The Model for predicting students academic performance provides secondary schools with an easy way of identifying areas which require attention in improving their outcomes as well as provide the school with the opportunity to harness data mining technology with the view of enhancing their curriculum. The system provides an easier method of determining students outcomes academically with the aid of technological advancements and also serve as a means where patterns can be recognized in the data gathered, if it is carried out genuinely and data analysis is properly carried out.

The prototype has been designed to meet both the functional and non-functional requirements stated in section three, so as to ensure the provision of a robust system that will sufficiently perform all tasks related to predicting students academic performance.

The model is also secure as it will be operated in confidentiality and access to student data will have to be authorized by specific individuals who can easily monitor and control the flow according to laid down procedures. Moreover, it takes someone with high competence skills in the field of ICT and with the knowledge on how to use the WEKA software to be able to manipulate the system for now.

5.2 Future Research

Further development of the model can be done in with enhancements incorporating more capabilities than the WEKA software since technology is dynamic and constantly changing. Also, it is possible to develop a software specifically for predicting students academic performance for use by our educational system citing the need to improve the outcome of students across various tiers of education as a basis. Other programs such as Python could be explored to develop different models in this regard as potentials in Artificial Intelligence are immense and diverse. In addition, a feedback system can be initiated with the aim of further improving on the requirements needed to carry out such task in future addressing the need of evolution in data mining and analysis. Furthermore, the contextual factors are not limited to the ones mentioned in this study. As such, other factors identified can be incorporated into improved models designed in the future.

6. REFERENCES

- [1] P. Cortez and A. Silva (2008). Using Data Mining to Predict Secondary School Student Performance
- [2] Abass,O., Oyekanlu ,E.A., Alaba, O.B,andLonge. O.B.(2011). Forecasting Student Academic Performance using Case-Base Reasoning. International Conference on ICT for Africa. 23-26th march, 2010 Otta Nigeria, pp.105- 112.
- [3] A.U. Mubarak (2019). International Journal of Computer Applications (0975 – 8887) Volume 178 – No. 48
- [4] Coker, Frank (2014). Pulse: Understanding the Vital Signs of Your Business (1st ed.). Bellevue, WA: Ambient Light Publishing. pp. 30, 39, 42, more. ISBN 978-0-9893086- 0-1.
- [5] Conz, Nathan (September 2, 2008). "Insurers Shift to Customer-focused Predictive Analytics Technologies", Insurance & Technology, archived from the original on July 22, 2012, retrieved July 2, 2012
- [6] Creswell, J. (2012). Educational research: Planning, conducting, and evaluating quantitative and qualitative research (4thed.)
- [7] Eckerson, Wayne (May 10, 2007), Extending the Value of Your Data Warehousing Investment, The Data Warehouse Institute
- [8] Fletcher, Heather (March 2, 2011). "The 7 Best Uses for Predictive Analytics in Multichannel Marketing", Target Marketing
- [9] Hashmia Hamsa, Simi Indiradevi, Jubilant J. Kizhakkethottam (2016). Student academic performance prediction model using decision tree and fuzzy genetic algorithm
- [10] J. L. Rastrollo-Guerrero, J. A. Gómez-Pulido and A. Durán-Domínguez (2020). Analyzing and Predicting Students' Performance by Means of Machine Learning: A Review
- [11] Kanakana, G.M.; Olanrewaju, A.O. (2001). Predicting student performance in engineering education using an artificial neural network at Tshwane university of technology.
- [12] McCulloch, Warren; Walter Pitts (1943). "A Logical Calculus of Ideas Immanent in Nervous Activity". Bulletin of Mathematical Biophysics.
- [13] Michael Dixon (2019). Data Analytics , Predictive Analytics
- [14] Neelam Tyagi (2022) Predictive Analytics: Techniques and Applications
- [15] Nyce, Charles (2007), Predictive Analytics White Paper(PDF), American Institute for Chartered Property Casualty Underwriters/Insurance Institute of America, p. 1
- [16] Oladokun, V. O., Charles-Owaba. O. E. & Adebajo, O. E. (2008). Predicting Student's Academic performance using artificial neural network: A case study of an engineering course. The pacific Journal of Science and Technology
- [17] Oyerinde, O. D., & Chia, P. A. (2017). Predicting Students' Academic Performances–A Learning Analytics Approach using Multiple Linear Regression. International Journal of Computer Applications, 975, 8887.
- [18] Oyerinde, O. D. (2019). Creating public value in information and communication technology: a learning analytics approach (Doctoral dissertation).
- [19] Rami Ali (2020). Predictive Modeling: Types, Benefits, and Algorithms
- [20] Reetesh Chandra (2018). Neural Networks: Applications in the Real World
- [21] Rosenblatt, F. (1958). "The Perceptron: A Probabilistic Model For Information Storage And Organization in the Brain". Psychological Review
- [22] Shaobo Huang and Ning Fang (2012). Early prediction of students' academic performance in an introductory engineering course through different mathematical modeling techniques
- [23] Smriti Dewan (2019). Top 5 Predictive Models and Their Applications
- [24] Sriram Parthasarathy (2021). Top 5 predictive analytics models
- [25] Tahmasebi; Hezarkhani (2012). "A hybrid neural networks-fuzzy logic-genetic algorithm for grade estimation". Computers & Geosciences