World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(RESEARCH ARTICLE)

Check for updates

Design and implementation of an expert system for predicting students' academic performance: A case study of computer science department federal polytechnic Bida

Godwin Chibuikem Chukwu [1,*], Monday Ernest Chukwu [2,*], Chidi Godwin Nweke [3], Paul Oluwafemi Oloye [4], Abiola Samuel Ajayi [5], Akorede Ayoola Asanre [6] and Olakunle Ibrahim Oresegun [7]

[1] *Federal Polytechnic Bida, Department of Information Communication Technology, Niger State, Nigeria.*
[2] *Department of Mathematics, Statistics and Computer Science, University of Calabar, Nigeria.*
[3] *Department of Computer Engineering, University of Ilorin, Kwara State, Nigeria.*
[4] *Department of Computer Science, Federal University of Technology Akure, Ondo State, Nigeria.*
[5] *Department of Computer Science, Texas Southern University, Houston, Texas, USA.*
[6] *Department of Mathematics, Tai Solarin University of Education, Ijebu Ode, Ogun State, Nigeria.*
[7] *Department of Mechanical Engineering, Zhejiang University Hangzhou, Hangzhou, Zhejiang, China.*

World Journal of Advanced Research and Reviews, 2024, 24(02), 915–926

Publication history: Received on 24 September 2024; revised on 09 November 2024; accepted on 11 November 2024

Article DOI: https://doi.org/10.30574/wjarr.2024.24.2.3459

Abstract

The application of machine learning techniques in predicting students' performance, based on their background and their in-term performance has proved to be a helpful tool for foreseeing poor and good performances in various levels of education. The major problem institutions face is the systematic monitoring of students' academic progress in their course of study, which is influence by many factors. This research work assesses the potential of predicting student academic performance with an expert system in the Federal Polytechnic Bida. The expert system is develop using python programming language and it accepts data about a student in various category such as demographic background, attitudes study habits, etc. and through the partial activation of multiple rules, predicts an outcome for each student. The students' data was collected via a user-friendly interface and analysed using fuzzy logic algorithm for prediction. The prediction analysis for students' can only be viewed by the administrator (instructor) as it requires logins for forecasting students' possible outcomes.

Keywords: Machine Learning; Student Performance Prediction; Academic Monitoring; Predictive Modeling; Fuzzy Logic Algorithm

1. Introduction

The importance of artificial intelligence in education has risen tremendously over the past few years; this is because it enables us to learn fresh, interesting, and practical information about things. Recent analysis has indicated that technology improvements have had a considerable impact on education, just as they have on many other facets of society and human endeavour [1]. The effectiveness of educational institutions is crucial in generating graduates and post-graduates of the highest caliber. By discovering instructional strategies that work for students from a variety of background information, student success prediction will aid educational institutions in improving learning and teaching practices [2].

One indication of course learning outcomes in a learning institution is excellent student performance. According to a summary of the literature by Yohannes & Ahmed [7] "Student success is defined as academic achievement, engagement in educationally purposeful activities, satisfaction, acquisition of desired knowledge, skills and competencies,

* Corresponding author: Monday Ernest Chukwu

persistence, attainment of educational outcomes, and post-college performance.". Although this is a multi-dimensional definition, authors in Molokomphale & Mhlauli, [8] gave an amended definition focusing on the most crucial six components, namely "Academic achievement, satisfaction, acquisition of skills and competencies, persistence, attainment of learning objectives, and career success." In universities, grade systems are used to assign a scale for evaluating students' academic performance, such as the Grade Point Average (GPA) or Cumulative Grade Point Average (CGPA). Grades are also a major component of academic achievement [3].

To boost students' academic performance and learning, educational institutions needed the outcomes of the predicate process. In order to increase students' academic success and learning, educational institutions need the predicate process' output [3]. The ability to predict student performance is a crucial area because it enables teachers to identify children who need further academic support [4]. A prediction model for student performance based on internal assessment was investigated by Hussain [4]. In this experiment, the maximum classification accuracy was generated by an artificial neural network (95.34%), however its potential to improve learning outcomes was not explored, leaving a conceptual vacuum. Numerous methods are being suggested right now to assess student achievement. Data mining is one of the most popular techniques for analyzing student performance, and it has recently gained significant traction in the field of education. In educational data mining, machine-learning approaches seek to model and identify significant hidden patterns and valuable information from educational environments.

Additionally, machine-learning techniques are used in the academic sector to represent a variety of student attributes as data points by applying them to big datasets. By fulfilling a variety of objectives, such as extracting patterns, anticipating behaviour, or finding trends, these strategies can be useful in a variety of sectors and help educators deliver the most effective teaching methodologies and track and monitor their students' development. Song & Lu [9] explored the use of machine learning to forecast student success in computer engineering using a decision tree algorithm. Similarly, Alyahyan & Düştegör [10] examined a machine learning-based model for predicting student performance in the context of Wolkite University where Support vector machines, neural networks, naive Bayes, and machine learning techniques algorithms were used. Renaningtias et al., [11] utilized computer learning to predict student performance; three different machine learning methods were used in his thesis (Naive Bayes classification, decision trees, and linear regression). The results show that student achievement may be correctly predicted using machine learning.

The usage of data mining techniques on educational data, such as student information, academic records, exam results, and student involvement in class, is referred to as educational data mining, or EDM [3]. An effective method for finding hidden patterns in educational data, forecasting academic success, and enhancing the learning and teaching environment is educational data mining, which is used to extract useful data and patterns from a huge educational database. Forecasting student performance using the essential data and patterns can assist teachers in creating effective teaching methods. The information gathered about instructional procedures provides fresh chances to enhance the educational process and maximize users' engagement with digital tools [5]. Improvements are made in a variety of areas as a result of the processing of educational data, including behaviour prediction for students, analytical learning, and novel approaches to education policies [6]. In addition to enabling educational authorities to build databased policies, this extensive data gathering will serve as the foundation for software designed using artificial intelligence to aid in the learning process. Yohannes & Ahmed [7] used data mining techniques to analyze student data and create a performance prediction model. Decision trees and other categorization techniques were employed in the study to assess student performance. The study also offered advice on how to support students as they learn and make timely decisions to prevent academic risk and desertion to help the impacted students perform better.

Predictive modelling is frequently use in educational data mining to forecast student achievement. Classification, regression, and categorization are three approaches used to develop prediction models. Classification is the most common job used to forecast pupils' success, several categorization techniques have been used to forecast student achievement, which includes Decision Tree, Artificial Neural Networks, Naive Bayes, Fuzzy Logic, K-Nearest Neighbour, and Support Vector Machine. To estimate students' academic performance for this study, data mining techniques was implored using fuzzy logic and decision tree was used to make the prediction. Ulloa-Cazarez [12]demonstrated how prediction techniques can be used on educational data using the fuzzy logic methodology to construct a model to forecast the performance of online students using grades collected during the first half of the scholar cycle. The study suggested integrating fuzzy logic techniques and genetic algorithms to improve the membership functions for future studies after discovering that fuzzy logic does not achieve a good accuracy level when compared with the statistical regression model.

Lekan [3] made use of support vector machines to predict student academic achievement by making use of the student's Grade Point Average. Their result showed that the Supportive vector machine performs better than other ML algorithms with an accuracy of 94% percent and 97% respectively. Renaningtias et al., [11] undertook a study to forecast student

performance using neural networks. In order to determine the applicability of neural networks, the study compared the prediction performance of neural networks against six different classifiers on this dataset. According to the study's results, neural networks perform better than other prediction models when generic training is used.

1.1. Statement of the problem

The academic success of students has a big influence on the generation of excellent graduates who will serve as remarkable leaders and human resources for the country, assisting in its economic and social development. As a result, need to design and implement an expert system for predicting students' academic performance. This will help identify students who have weaknesses in their course of study so that special attention can be given to them to help improve their performance and produce quality graduates.

Aim and objectives of the study

The objectives of the study are:

- To identify some appropriate variables that influence students' performance.
- To design and implement a prediction system
- To predict the academic performance of student in computer science department.

2. Methodology

2.1. Description and analysis of previous system

Currently, there are many prediction systems existing in various institutions in Nigeria. An example is a system adopted by the Department of Computer and Information Sciences, Tai Solarin University of Education, Ogun State Nigeria in which prediction is done using a neural network to establish and analyze the complex nonlinear relationship that exists between cognitive and psychological variables that influence the academic performance of secondary school students. In spite of the high level of prediction accuracy of ANN in nonlinear phenomena, however, the model does not easily allow the identification of how predictor variables are related to one another in the explanation of the academic outcome. In other words, ANN the model does not specify an explicit mathematical model for the relationship between inputs and outputs, hence the need for further research in this regard.

2.2. Description and Analysis of the Current System

In the Federal Polytechnic Bida, prediction of students' academic performance has been done in various forms, but all the methods still boil down to the manual process, for instance in the exam's office, students' academic performance is easily predicted based on their performance on various courses, lecturers, on the other hand predict students' performance based on responsiveness during lectures, coordinate, on and seriousness, though most times this methods involves generalizing the student performance using a key a factor.

2.3. SHORTCOMNGS OF the Current System

Despite the use of various manual forms of prediction by the department, authenticity can still not yet be verified, as most times students' performance is generalized and not done individually which in turn is not justifiable.

The current system is limited to:

- Prediction is done based on outcomes of exams and student position in class
- Factors influencing the performance are not determined
- There is usually variation in prediction based on different individual perspectives.

2.4. data collection method for the new system

To develop the model, a number of steps was followed which include: data collection, data processing, analysis, code before the prediction system will be realized. The questionnaire was created based on the study, and professionals from a variety of professions filled it out. There are two components to the questionnaire: (Personal Factors and Academic Factors). Various specialists will receive the questionnaire for analysis. In order to create a prediction model, the data are analyzed. Tables 1 and 2 display the questionnaire for the students' personal and academic data that was given to different specialists for analysis.

2.5. Description and analysis of the new system

Predicting a system is usually done by learning from the past for which historical data is obtained and analysed to study the results. The new system will have undergone data training as the questionnaires distributed were analysed by experts on areas and factors that will be used for prediction. There are numerous benefits of the proposed new system, this includes:

- Prediction accuracy will increase since analysis was done by experts in the selected fields
- The system analysis each variable separately before drawing conclusion
- Effect of each individual variable can be determined e.g., academic factors
- Students can be assisted in individual areas with emphasis
- Prediction is fast and reliable
- System can easily be modified and updated

3. Results

3.1. System design

System design is the process of defining the components, modules, interfaces, and data for a system to satisfy specified requirements. System development is the process of creating or altering systems, along with the processes, practices, models, and methodologies used to develop them. For the purpose of the system UML diagrams is used to explain the system and its functionalities. A UML diagram is a diagram based on the UML (Unified Modelling Language) with the purpose of visually representing a system along with its main actors, roles, actions, artifacts or classes, in order to better understand, alter, maintain, or document information about the system.

3.1.1. Flowchart Diagram

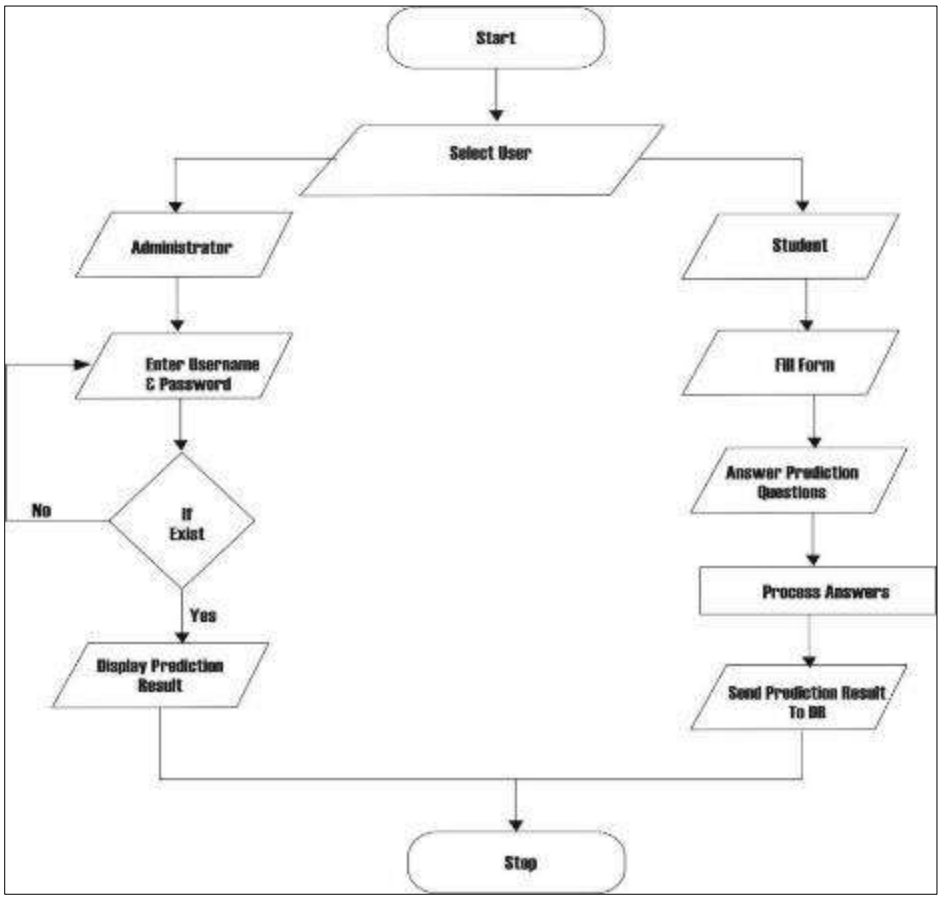

Figure 1 Flow Chart of the Student Academic Performance Prediction System

A flowchart is simply a graphical representation of steps. It shows steps in sequential order and is widely used in presenting the flow of algorithms, workflow or processes. Typically, a flowchart shows the steps as boxes of various kinds, and their order by connecting them with arrows. Figure 1 below shows the flow chart of the student academic performance prediction system. The flow chart comprises of the following:

- **SELECT USER:** The system gives room to select the type of users (Administrator or Student)
- **ADMINISTRATOR:** This user is also known as the instructor; the system asks for login credentials before access is given to this user.
- **STUDENT:** Every student is requested to fill a form; this form contains their personal data before they can be allowed to make prediction.
- **ANSWER PREDICTION QUESTION:** The system prompts series of questions for the student user, these questions are analysed following the selected algorithm, before prediction is done for each student based on their response to the questions prompted by the system.

3.1.2. Use Case Diagram

A use case diagram at its simplest is a representation of a user's interaction with the system that shows the relationship between the user and the different use cases in which the user is involved. A use case diagram can identify the different types of users of a system and the different use cases and will often be accompanied by other types of diagrams as well. While a use case itself might drill into a lot of detail about every possibility, a use case diagram can help provide a higher-level view of the system. It has been said before that "Use case diagrams are the blueprints for your system". They provide the simplified and graphical representation of what the system must actually do.

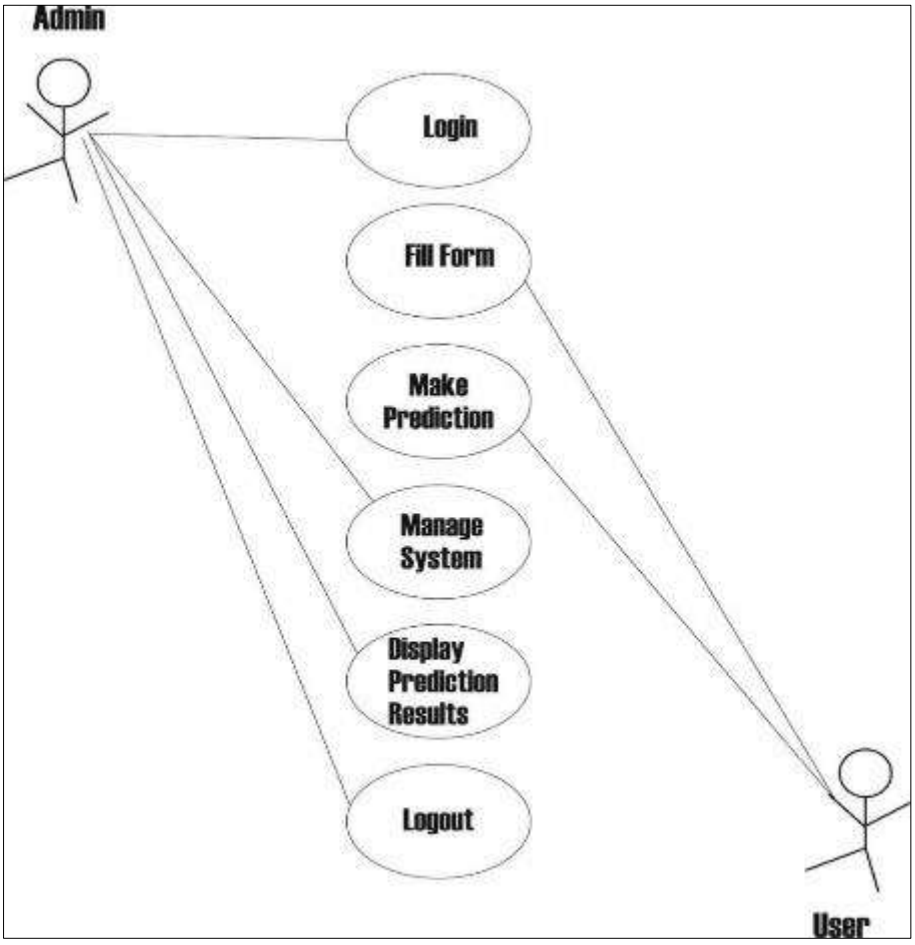

Figure 2 Use case diagram of the student's academic performance prediction system.

The use case diagram for this prediction system as depicted in Figure 2 above consists of two actors - the administrator and the student. The system also consists of six use cases – two of which are operated by user while the administrative officer operates on the rest of the cases. The activities of the student include fill a form, and make prediction. While the

activities of the administrator comprise of the activities of the user (student) together with other activities such as view/ display prediction of students and management of the system.

3.1.3. Data Flow Diagram

A dataflow diagram is a graphical representation of the "flow" of data through an information system, modelling its process aspects. A DFD is often used as a preliminary step to create an overview of the system without going into detail, which can later be elaborated. DFDs can also be used for the visualization of data processing. A DFD shows what kind of information will be input to and output from the system, how the data will advance through the system, and where the data will be stored.

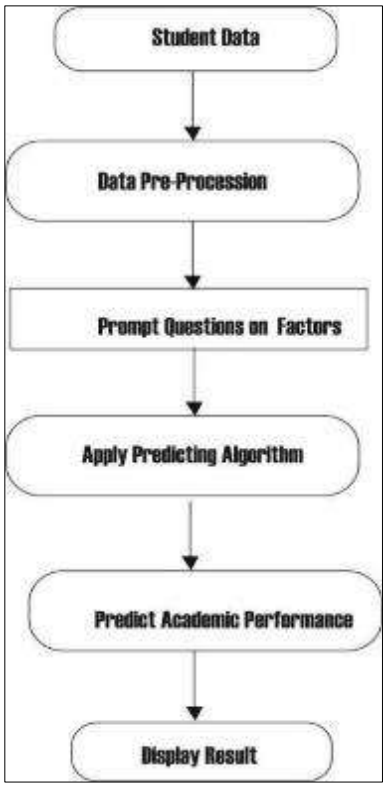

Figure 3 Represents the data flow between datasets, user and prediction model

The data flow diagram of the Figure 3 above shows the movement of data. After successfully filling the form, questions are then prompted for the user to answer based on some of the selected factors that affects students' academic performance. After which the system applies the algorithm for prediction and stores the user (student) prediction result in the database of the system.

3.2. Input specification and design

3.2.1. User Selection

Table 1 User Selection

Field name	Data type(value)	Constraint
Administrator	Button	Not null
User	Button	Not null

Figure 4 below shows the system user selection interface, the system comprises of two users an administrator (lecturer / instructor) and a user (student) each is expected to click on the button that designated for their level.

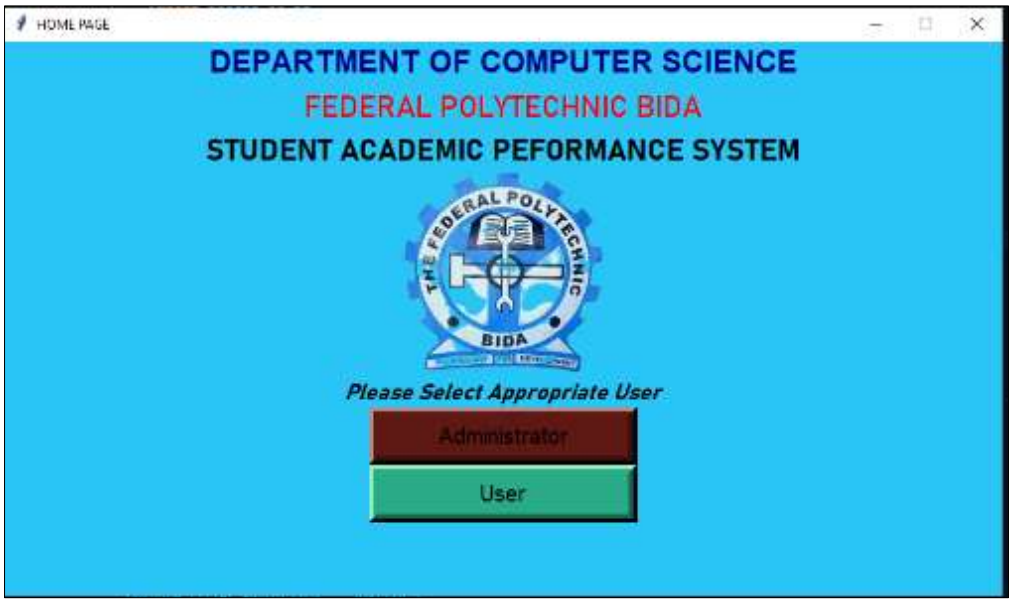

Figure 4 User Selection Interface

3.2.2. Administrator Form Login

Table 2 Administrator login

Field name	Data type(value)	Constraint
Username	VarChar (10)	Primary key
Password	VarChar (10)	Not null

The administrator Login Page as depicted in Figure 5 below provides the administrator with a GUI to login the software, it requires the administrator to enter a username and password.

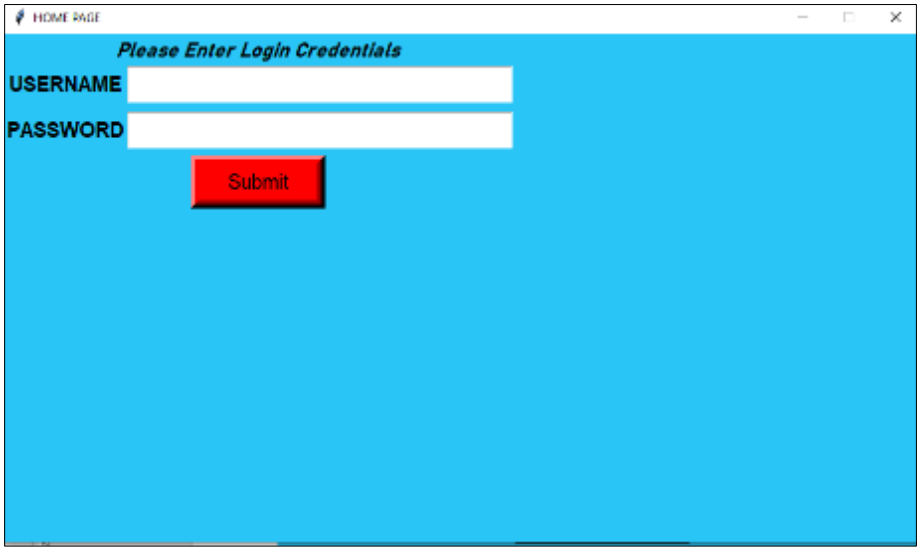

Figure 5 Administrator Login Page

3.2.3. User Registration Page

Table 3 User Registration Page

Field name	Data type(value)	Constraint
Full name	VarChar(30)	Not Null
Matric No	VarChar(10)	primary key
Department	VarChar (15)	Not Null
Semester	VarChar(7)	Not Null
Email	VarChar (20)	Not null
Telephone	VarChar(11)	Not Null
Address	VarChar (30)	Not null

The user registration page as depicted in Figure 6 is a form that pop-up whenever the button “Student” is selected in the user selection page. It is a form that a user is expected to fill before prediction can take place.

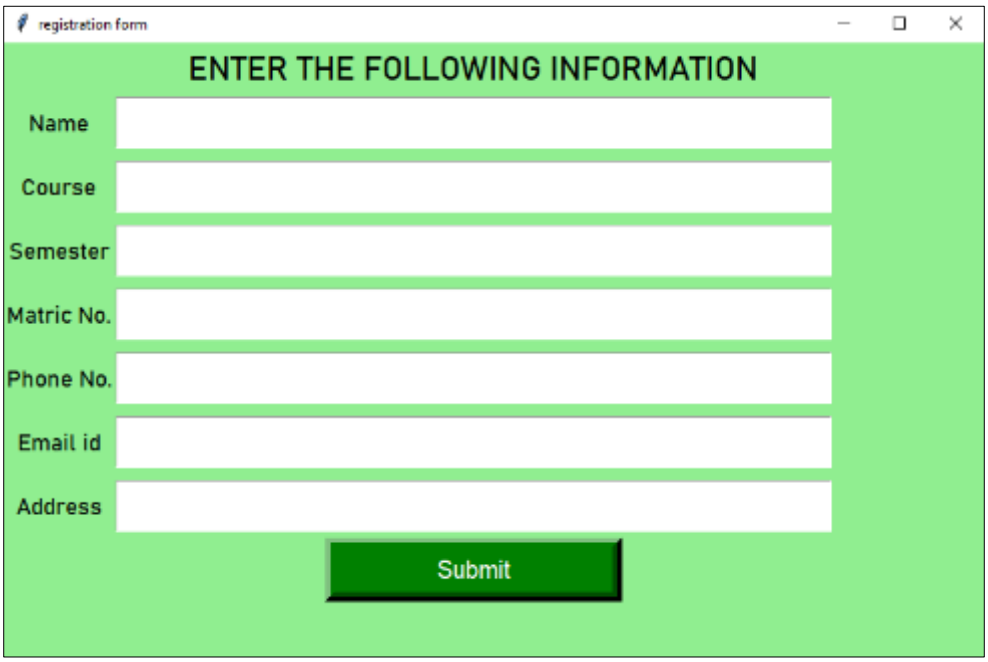

Figure 6 User Registration Page

3.3. Prediction Page

The prediction page as depicted in Figure 7 below is an interface which comprises of series of questions which are in Radio Button format.

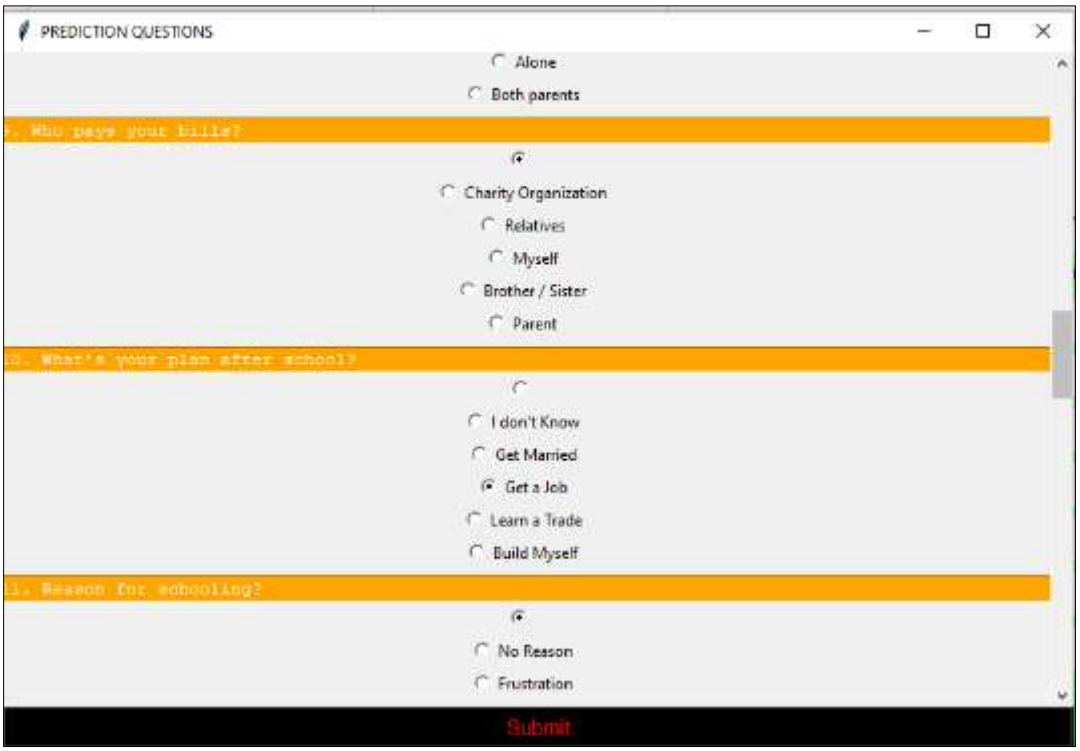

Figure 7 Prediction Page

3.4. Output specification and design

3.4.1. Invalid Login Credential

The invalid login credentials as depicted in Figure 8 below display whenever the admin login credentials entered does not match with that set by the system management or when a user enters wrong login credentials.

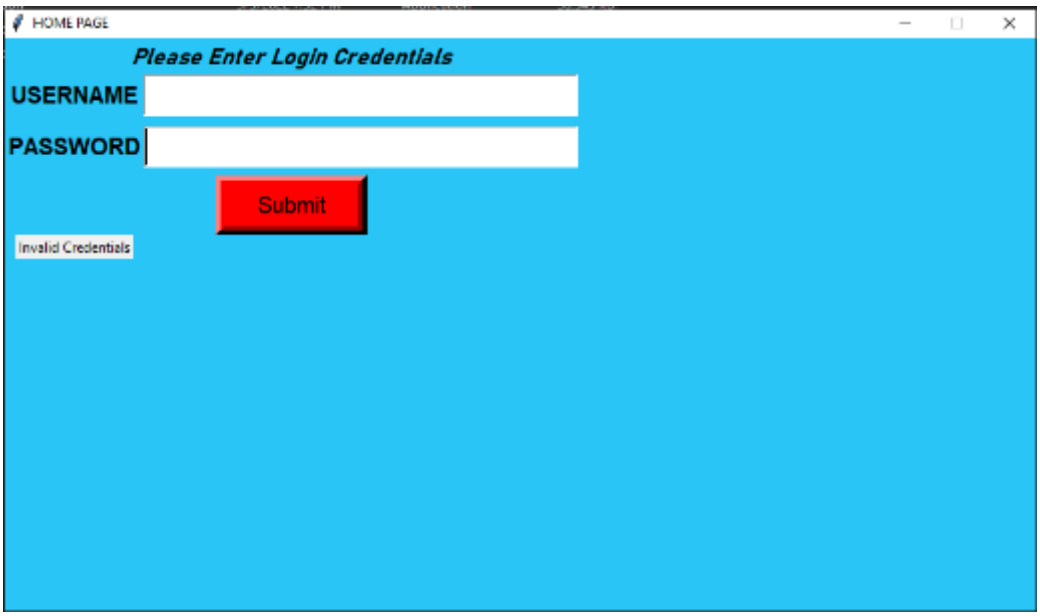

Figure 8 Invalid Login Credentials

3.4.2. Response Submitted

The response submitted as depicted in Figure 9 below is a message box whenever the user (student) has finished making prediction (answering prediction questions) it is a confirmation message showing the users answers has been submitted.

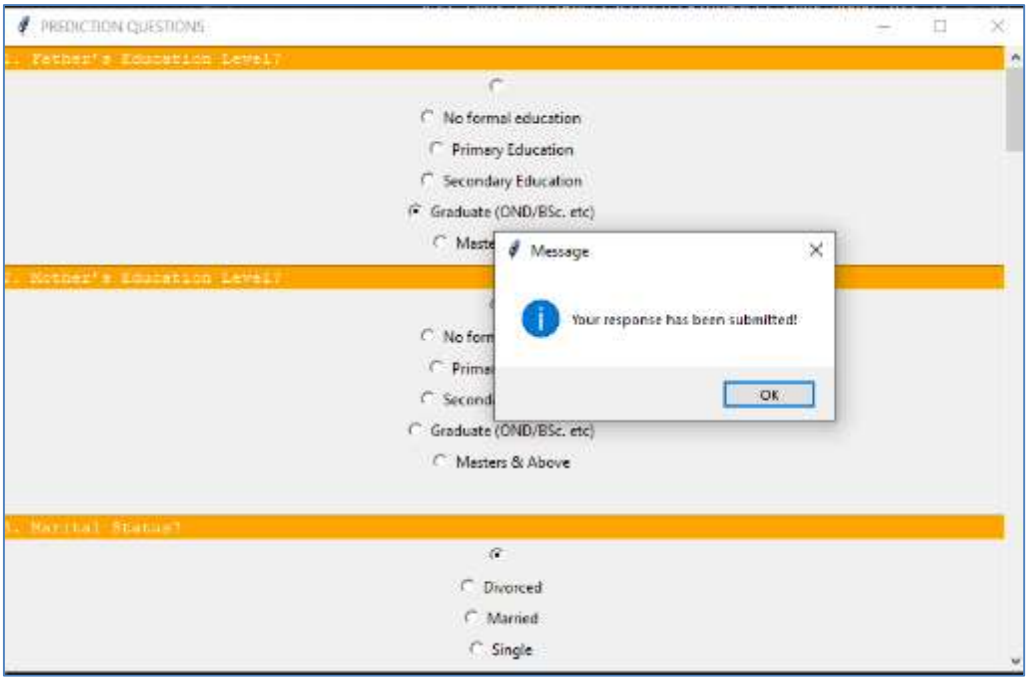

Figure 9 Response Submitted

3.4.3. Display Prediction

Display Prediction as depicted in Figure 10 below is use by the administrator view list of users that (students) that have made prediction and the outcome of their prediction analysis.

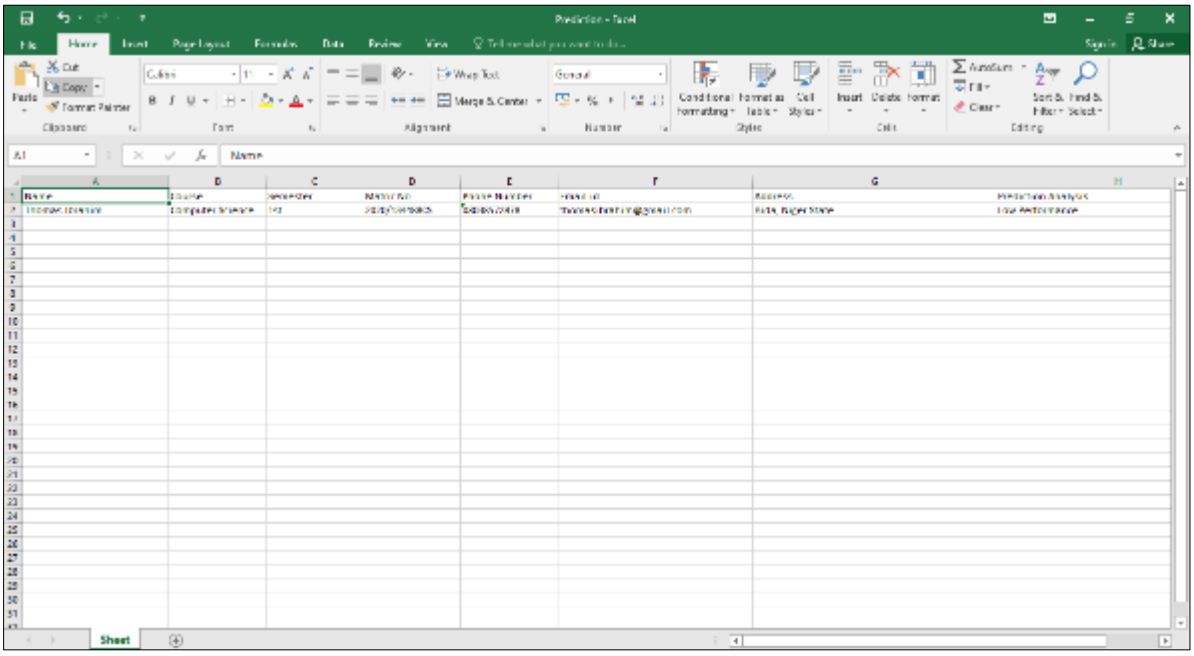

Figure 10 Prediction Record stored in database

4. Conclusion

The ability of the educational institution to provide the student with additional support is facilitated by the ability to predict a student's performance in advance. For at-risk students who have trouble getting good scores in the class, an early performance prediction would be helpful. It's crucial to predict these students' performance on a regular basis so that they can receive the support they need to learn and advance. According to research, it is possible to anticipate a student's academic performance by looking at the things that influence them. When the forecast is accurate, the teacher may spend some extra time helping pupils who have poor academic performance raise their final examination scores because they are considered to be at risk.

Recommendations

Prediction of students' academic performance plays a vital role in institutions, as it tells lectures students possible result, thereby giving room for necessary assistants to students. It is therefore recommended that the prediction system be implemented in the computer science department, to enable lecturers provide the necessary assistance to students with low performance.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

Statement of informed consent

Informed consent was obtained from all individual participants included in the study.

References

[1] Chen X, Xie H, Hwang GJ. A multi-perspective study on artificial intelligence in education: Grants, conferences, journals, software tools, institutions, and researchers. Comput Educ Artif Intell. 2020;1. doi:10.1016/j.caeai.2020.100005.

[2] Nosseir A, Fathy YM. A mobile application for early prediction of student performance using fuzzy logic and artificial neural networks. Int J Interact Mob Technol. 2020;14(2):4–18. doi:10.3991/ijim.v14i02.10940.

[3] Lekan AJ, Oloruntoba SA, Akinode JL. Student academic performance prediction using support vector machine design and development of automated lecture-time-tabling system view project computer security and networks view project ijesrt international journal of engineering sciences & research technology student academic performance prediction using support vector machine. Int J Eng Sci Res Technol. 2019;588. doi:10.5281/zenodo.1130905.

[4] Hussain S, Muhsin ZF, Salal YK, Theodorou P, Kurtoğlu F, Hazarika GC. Prediction model on student performance based on internal assessment using deep learning. Int J Emerg Technol Learn. 2019;14(8):4–22. doi:10.3991/ijet.v14i08.10001.

[5] Zhang Y, Yun Y, An R, Cui J, Dai H, Shang X. Educational data mining techniques for student performance prediction: Method review and comparison analysis. Front Psychol. 2021;12. doi:10.3389/fpsyg.2021.698490.

[6] Khalaf A, Majeed A, Akeel W, Salah A. Students' success prediction based on Bayes algorithms. Int J Comput Appl. 2017;178(7):6–12. doi:10.5120/ijca2017915506.

[7] Alyahyan E, Düştegör D. Predicting academic success in higher education: literature review and best practices. Int J Educ Technol High Educ. 2020;17(1). https://doi.org/10.1186/s41239-020-0177-7

[8] Molokomphale L, Mhlauli MB. An investigation on students academic performance for junior secondary schools in Botswana. Int J Environ Sci Educ. 2014;3. http://www.eurojedu.com

[9] Ulloa-Cazarez RL. Fuzzy logic model for predicting student performance on online courses. 2014. https://doi.org/10.13140/RG.2.1.1236.6720/1

[10] Yohannes E, Ahmed S. Prediction of student academic performance using neural network, linear regression and support vector regression: a case study. Int J Comput Appl. 2018;180(40):39–47. https://doi.org/10.5120/ijca2018917057

[11] Song YY, Lu Y. Decision tree methods: applications for classification and prediction. Shanghai Arch Psychiatry. 2015;27(2):130–5. https://doi.org/10.11919/j.issn.1002-0829.215044

[12] Renaningtias N, Suseno JE, Gernowo R. Hybrid decision tree and naïve bayes classifier for predicting study period and predicate of student's graduation. Int J Comput Appl. 2018;180(49).