Web-Based Grade Prediction System

Izang Aaron. A[a*]., Abdullah Halimah. D[b]., John Henry. C[c]., Oyebode Oluwajenyo. B[d]., & Wami Praise. M[e]

[a]*Department of Information Technology, School of Computing and Engineering Sciences, Babcock University, Ilishan-Remo, Ogun State, Nigeria.*

[b,d,e]*Department of Computer Science, School of Computing and Engineering Sciences, Babcock University, Ilishan-Remo, Ogun State, Nigeria.*

[c]*M.A. Rare Books and Digital Humanities, UFR- Sciences du langage, de I'l homme et de la société. University of Franche-Comté, Besançon, France.*

[a]Email; izanga@babcock.edu.ng
[b]Email; Dalimat2000@gmail.com
[c]Email; henry.john@edu.univ-fcomte.fr
[d]Email; jenyooyebode@icloud.com
[e]Email; manuchimwami@gmail.com

ABSTRACT

The need for improving the educational system has led to researches in the area of educational data mining, which involves the process of applying data mining tools and techniques to analyse data at educational institutions. In some higher institutes, students are plagued with the problem of having to struggle hard to complete different courses since there is no dedicated support offered to students who may need special attention in some of the registered courses. These problems arose due to lack of system to analyse and monitor students' progress and performance in their course at interval. This has therefore resulted in poor academic performance and sub-optimal aachievement by students. To this end, this study developed a system that helps to track students' assessments and predict their grade for different courses. The system was fed with the students' data such as the attendance in class, and number of hours puts in for studies in each course by the student. This two are now used as a criterion to predict the performance in a particular course by getting the average of both. The system was developed using PHP for the backend, MySQL for the Database and HTML, CSS and JavaScript for the front end. Students' data obtained were modelled to predict student grades in their related courses, and levels using a coded algorithm in PHP.

Furthermore, the outcome of this research can be used by lecturers who are major module in the system, in higher institutions to track their students' academic performance and help in advising student in subjects they need to improve.

1. INTRODUCTION

Student's academic performance and grades are affected by many factors which include personal, socio-economic and other environmental variables [1]. Knowledge about these factors and their toll on student's grades can help in managing their effects. Recently, much attention has been paid to educational mining research. Educational Data Mining refers to techniques, tools, and research designed for automatically extracting meaningful data from large repositories of data generated by or related to people learning activities in educational environment [2]. Predicting student's grades or performance has become more challenging due to the large volume of data in educational databases.

In this modern era of information technology, the repository of students' result contains a lot of data. The student repository contains scores and grades of courses for which students are enrolled in the institution. This data bank can be used for knowledge mining; however, many educational institutions around the world and in Nigeria do not apply any knowledge unearthing processes to these student's data banks. This knowledge discovery process can be applied to improve the students' grades in their respective courses, enhance student teacher interaction, identify interdependency among courses, uncover student failure pattern, predict student final result, and to improve the overall quality of education. The ability to predict student performance is very important in educational environments. Increasing student success is a long-term goal in all academic institutions [3].

If educational institutions can predict students' grades early before their final examination, then extra effort can be taken to arrange proper support for the low performing students in that particular course, so as to improve their studies and help them to succeed. On the other hand, identifying attributes that affect course success rate can assist in courses improvement. Decision tree is a commonly used tool in data mining that utilizes tree structure to build regression or classification models. Wherein the dataset is divided into smaller subsets and at the same time an associated decision tree is incrementally developed [4].

This study ffocuses on developing a system that would predict student grades in various courses using an algorithm. This will identify the area of weakness and strengths of the student in their registered course, as well as identify the student success rate, then use these factors as an early predictor for expected success rates and handle ttheir weaknesses.

1.1 GRADE PREDICTION DEFINED

Grade prediction is simply the act of forecasting a grade. A predicted grade is essentially the qualification or score that students' school or college believes they are likely to obtain, given a number of circumstances [5]. These grades that are predicted are then used by the college or university to understand the potential of a student, and whether or not they might require assistance in their academics. The grade prediction process can bring up challenges such as:

i. Changes in academic and technical qualifications; that is, changes to content, assessment, etc.
ii. Differences in Performance with Evidence when from peers, parents and the media.
iii. External and internal pressures
iv. Divergences in university standards

However, considering all of that information, it is still a given that predicted grades should be recorded regardless of all pending qualifications, unless the method of assessment for the qualification structure is considered inappropriate. It should be in the best interest of students because success is the end goal. It should also be aspirational, but achievable. Stretching a predicted goal can be motivational for a student. They should also be determined

by professional judgement because expertise and experience are very vital in determining predictions [6]. Most importantly, predicted grades should be driven by data provided by the university, specifically the department where a student is registered. Each school has its own process. But ideally, they should consider the second and third levels of performance.

Predicted grades should not be affected directly by a student, parent, guardian or care taker's pressure. There are several risks associated with inflated and suppressed grades that are predicted. They should not be affected by behavior or background either. Rather, they should be made objectively and disregard external factors.

2. RELATED WORKS

The ability to predict the students' academic performance is very important in institution educational systems, as well as for faculty, university and educators and for students. Data mining in the field of education (Educational Data Mining - EDM), as a new field of research, has developed in the last decade as a special area of application techniques and tools for detecting regularities and correlations in the data (data mining), with the aim of analyzing the unique data types that appear in educational system for solving various problems of educational and instructional improvement process [7]. EDM is engaged in development, research and application of methods to detect data regularities in the database in the field of education, which would otherwise be difficult or almost impossible to analyze and determine the dependency patterns of behavior and learning among students, primarily because of the large amount of data.

Applications of data mining techniques in educational environments are more popular in areas of research and there is an increasing number of research papers published in the last years [8]. Researches in educational data mining areas focus on different aspects of educational processes: students, teachers, teaching aids, organization of teaching etc. [9] One of the main goals and the basic application fields of educational data mining can be classified into the following categories: Prediction of students' success, Organization of teaching programs, Prediction of enrolment of students to the higher level of educational program, Discovering cheating during an online examination, Identification of abnormal/extreme values in the educational system. In order to achieve these goals, different algorithms of data mining are used, such as: Decision trees, Artificial Neural Networks (ANNs), K-Nearest Neighbor, Naive Bayes, Support Vector Machines (SVM), Cluster algorithms, etc.

Traditionally, academic researchers have used statistic models and methods in order to predict the success of the students. Today, there are many different approaches about classifying the students and predicting their grades. Researchers tried to predict the success of the students, using the linear regression and neural network models. The results of their research point to the fact that the models of logistic regression do not predict the behavior of students equally as well as the artificial neural network models. They also compared the performance of the three different techniques of data mining for predicting students' behavior: neural networks, cluster algorithms and decision trees, where the model based on neural network gave the best results [10]. Furthermore, they used neural networks to predict the success of the students on the exams, defined with binary classes (pass or fail). [11] compared two data mining techniques, ANN and the combination of clustering and decision tree classification techniques for predicting and classifying students' academic performance, and based on that research they identified the patterns that influence or affect the student's academic performance.

[12] presented two prediction models for the estimation of students' performance in final examination. The work made use of the popular dataset provided by the University of Minho in Portugal, which relates to the performance in math subject and it consists of 395 data samples. Forecasting the performance of students can be useful in taking early

precautions, instant actions, or selecting a student that is fit for a certain task. The need to explore better models to achieve better performance cannot be overemphasized. Most of earlier work on the same dataset used K-Nearest Neighbor algorithm and achieved low results, while Support Vector Machine algorithm was rarely used, which happens to be a very popular and powerful prediction technique. To ensure better comparison, Support Vector Machine algorithm and K-Nearest Neighbor algorithm were both applied on the dataset to predict the students' grades and then compared their accuracy. Empirical studies outcome indicated that Support Vector Machine achieved slightly better results with a correlation coefficient of 0.96, while the K-Nearest Neighbor achieved a correlation coefficient of 0.95.

[13] Here, researchers employed methods for supervised machine learning. Data gathered from 489 students in the department of information and communication technology at the northwest Malaysia Polytechnic throughout the course of the four preceding academic years, from 2016 to 2019, were used in the research. To compare the effectiveness of classification and regression techniques in predicting students' final grades, various tests were conducted utilizing Decision Tree (J48), Random Forest (RF), Support Vector Machines (SVM), and Logistic Regression (LR). According to the results, J48 was the best predictive analytics model with the greatest prediction accuracy rate of 99.6%, which might help identify student dropouts early on and help educators maintain their position as excellent performers in higher education.

[14] The study employs supervised ML methods to forecast students' grades and test scores. The Board of Intermediate & Secondary Education in Peshawar, Khyber Pakhtunkhwa provided the data set for this study (B.I.S.E.). The proposed technique involves pre-processing the acquired data set to assure data quality and training the regression model and DT-classifier with labelled academic historical data of students (composed of 30 ideal attributes). The output of the models will then be looked at. While the classification method will forecast marks, regression will forecast grades.

[15] Researchers used two trials to compare the performance of a K-NN method and a multiple linear regression technique. In terms of predicting the final score, the multiple linear regression model performed better than the K-NN classification model. The linear regression model performed better than the baseline, while the K-NN classification model did not. The findings of the multiple linear regression showed that the number of words used, the number of spelling and grammar errors, and the reading grade level were all significant predictors of the final score.

[16] In order to predict undergraduate students' final exam scores, the study proposes a new machine learning-based model that uses the findings of the midterm exam grades as the main data. The performances of the machine learning methods random forests, nearest neighbour, support vector machines, logistic regression, Naive Bayes, and k-nearest neighbour were calculated and compared to predict the students' final exam scores. The academic performance grades of 1854 students who took Turkish Language-I at a state university in Turkey during the autumn semester of 2019–2020 made up the dataset. The findings indicate that the suggested model has a classification accuracy of between 70 and 75 percent. Only three different types of parameters—midterm exam grades, departmental data, and faculty data—were used to make the forecasts.

[17] Comprehensively evaluated real-world data gathered from the Information Technology University (ITU), Lahore, Pakistan, using the techniques of collaborative filtering (CF), matrix factorization (MF), and Restricted Boltzmann machines (RBM). They assess the academic standing of ITU students who were accepted into the Electrical Engineering department's bachelor's program. The RBM method is discovered to be superior to the other methods used to forecast students' success in a certain course. The efficacy of the employed RBM technique has been empirically validated on real-world datasets. By predicting how

well other students in the course performed, the feedback model technique allows us to assess the students' understanding of a specific course area and then offer them with the necessary guidance about other courses in that domain.

3. METHODOLOGY

The software development model that was used is the incremental model. This model was used because it allows changes and additional functionality at a later time which makes development fast and more features would be easy to add.

XAMPP package was used for the development of the entire system, Apache web server was used as the sever to host the system locally, MySQL was used to create the database that houses all the data and information entered into the system by the administrator and the other staff members with access to the system using PHPMyAdmin. HTML, CSS were used to design the web pages for the application, and PHP was used for building the backend of the application. Figure 1 shows the use case diagram of the web-based grade prediction system.

Figure 1. Use case diagram of the system showing the actors and their functions

Use Case:

The actor in this use case as shown in figure 1 is the staff and he is able to perform actions which include:

i. Login: The login page gives access to the staff members registered in the system. This action allows them to have full access to the pages allowed by the administrator on the system.

ii. Add/edit data: The system allows the user to enter and alter data to the system. The data that is entered is what will eventually be used in the prediction of grades by the system.

iii. Make a prediction: using all the data registered in the system, the user can predict the grade of a student and even view all the other data recorded that belongs to the student, as is entered in the database.

iv. Save predictions: after a prediction has been made, the user can save the prediction in the database and view it anytime he needs to

    v.   Exporting predictions: the user can also go as far as exporting predicted grades to use at a later time for other requirements he might have.

    vi.   Add suggestions: Users also have the ability to make suggestions to the administrator or management board of the system

There are two main actors in the figure 1 above, which are the user and the administrator. Their functions are as follows:

    i.   **Users:**

          These are the staff members that will use the system or the platform. They will be granted access by the administrator and be able to use the functions that the system provides, and they will be able to do the following:

    a. Login
    b. Add and edit data
    c. Make predictions with the system
    d. Export the predicted output
    e. Provide feedback for the administrator
    f. Logout

    ii.   **Administrator**

          This actor monitors all predictions and events happening on the platform. He is registered by default and can create users who will enter and use data in the system. He is also able to do the following:

    i.   Create new users
    ii.   Login
   iii.   Add and edit data
   iv.   Review feedback
    v.   Logout
   vi.   Manage the system

3.1   WORKFLOW PROCESS OF THE WEB-BASED GRADE PREDICTION SYSTEM

The Web-Based Grade Prediction system is designed to work following a step-by-step process as described below

    i.   Use student matric number to get all registered courses of the student
    ii.   Get course attendance and study time score requirement for selected course
   iii.   Get average course attendance and study time score history of matric no
   iv.   Study time and attendance rating of student is always updated when new course has been graded
    v.   Apply mean average of study time and attendance rating of student against the required rating of the course to get a percentage value
   vi.   Add mean average of study time and attendance score to get a percentage value (which is over 100%)
   vii.   The score generated is displayed for upward or downward review by the lecturer who generated the score on the staff module.

The above entire process was implemented using a function in PHP to ensure the accuracy of the grade prediction is optimal.

Figure 2 shows a diagram that describes the workflow process of the web-based grade prediction system.
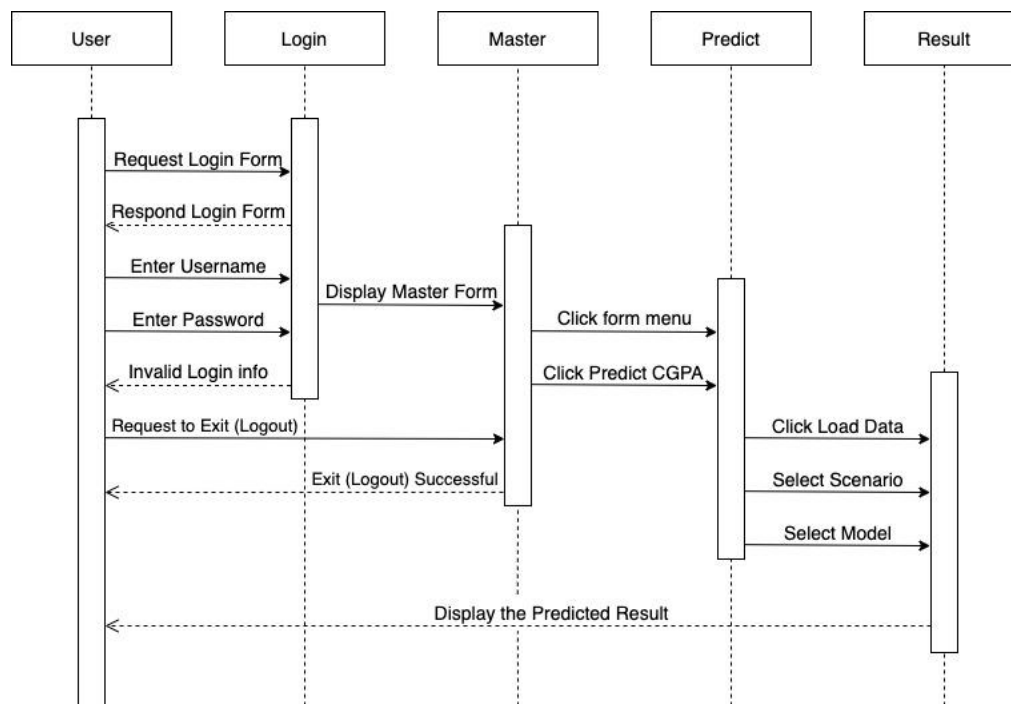
Figure 2. System Sequence Diagram

The diagram above explains a workflow of the sequence of the grade prediction system, which is also the backend of the system that specifies the operation of this system.

4. IMPLEMENTATION OF THE WEB-BASED GRADE PREDICTION SYSTEM

The system designed comprises of the front-end and the back-end. The front-end of the system includes the interface of the system, such as, what the users can see. The front-end components include the various pages of the system which includes the Dashboard, Administrator's module page, the datasets page, the registered staff page. Furthermore, these pages (front-end) are linked to the back-end component which contains the database of the system. Using PHP scripting language, the database was connected to the front-end.

4.1 The Dashboard

A Dashboard is normally the primary web page a tourist navigating to a website from an internet browser will see, and it could also function as a landing page to draw visitors. The dashboard is used to facilitate navigation to other pages on the website. Figure 3 shows the dashboard of the web-based grade prediction system.
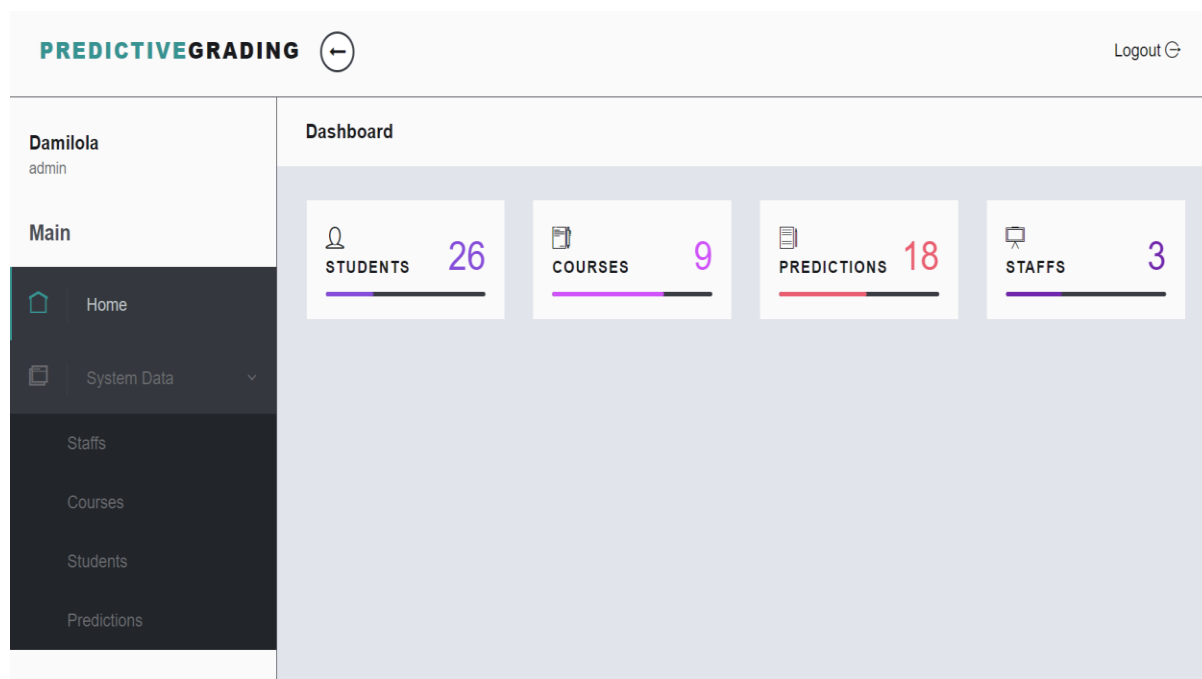
Figure 3. The dashboard of the grade prediction system

4.2 Administrator's Module

An administrator is a special user who has the authority to make changes on the software and effects that will affect customers as an update when they login. Several or various administrators can be employed in the system with the task of login, manage users, etc. figure 4 shows the admin module of the system
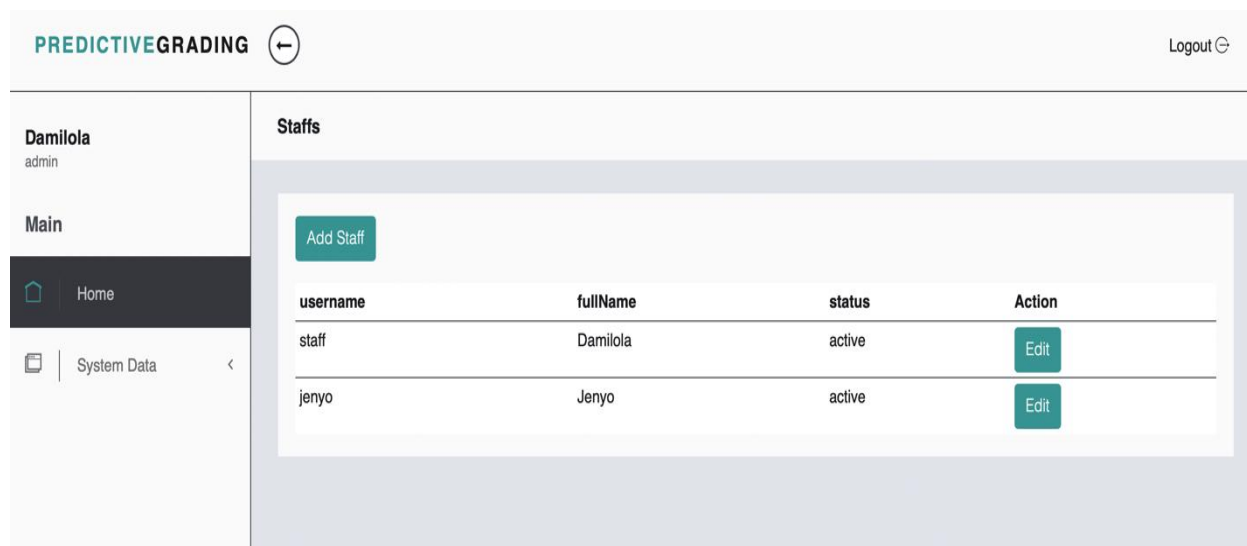

Figure 4. The administrator's module page showing managed staff

4.3 Staff Module

The staff are the individuals for which the system is designed. They are the users of the application, and since this is a predictive grading system, the intended users are the staff who will be grading students total score for a specific course or number of courses.

The module provides the staff with the interface that enables the user to access their dashboard to perform activities that lead to predicting student course scores. The staff can

add students and make prediction based on the students' academic history. Figure 5 shows the   staff   module   and   the   various   task   they   can perform.
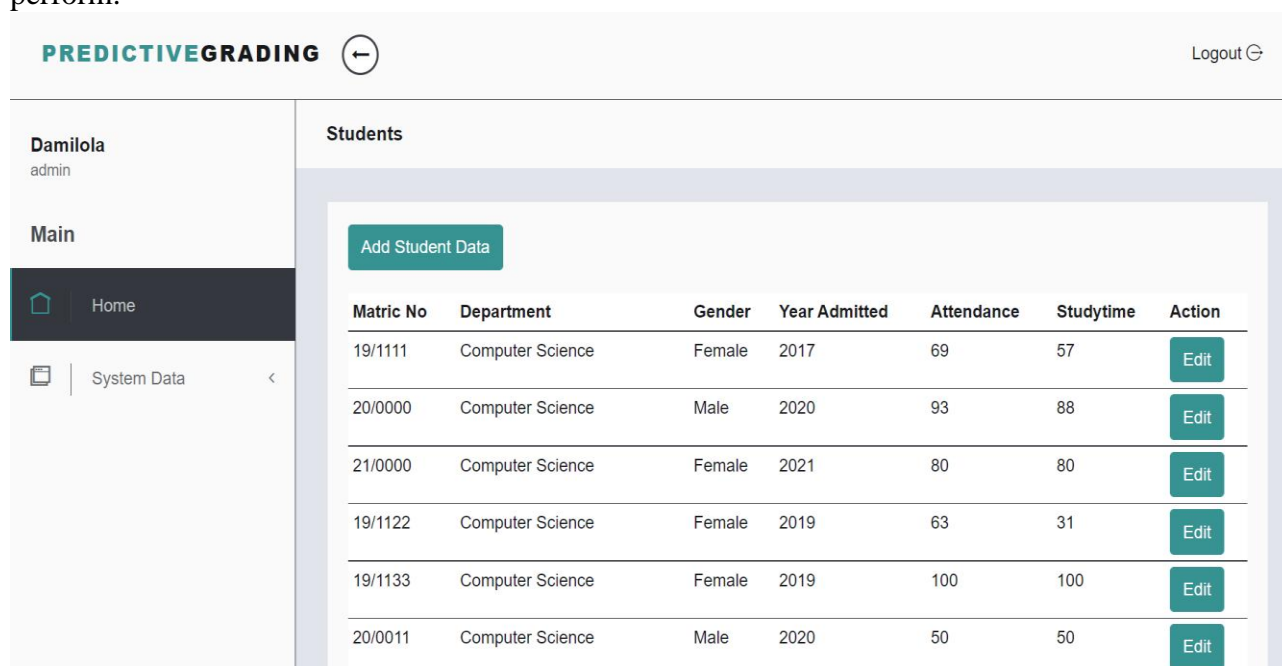

Figure 5. The staff module page showing registered data

4.4   The Prediction Page

The staff can begin the prediction process for a student based on history of performance on courses the student has taken, which are already registered in the database. The staff enters the student matric number and then clicks on the start button. They then go on to follow all the other procedures as directed by the system until they arrive at the desired prediction for the student. Figure 6 shows the prediction page of the system.
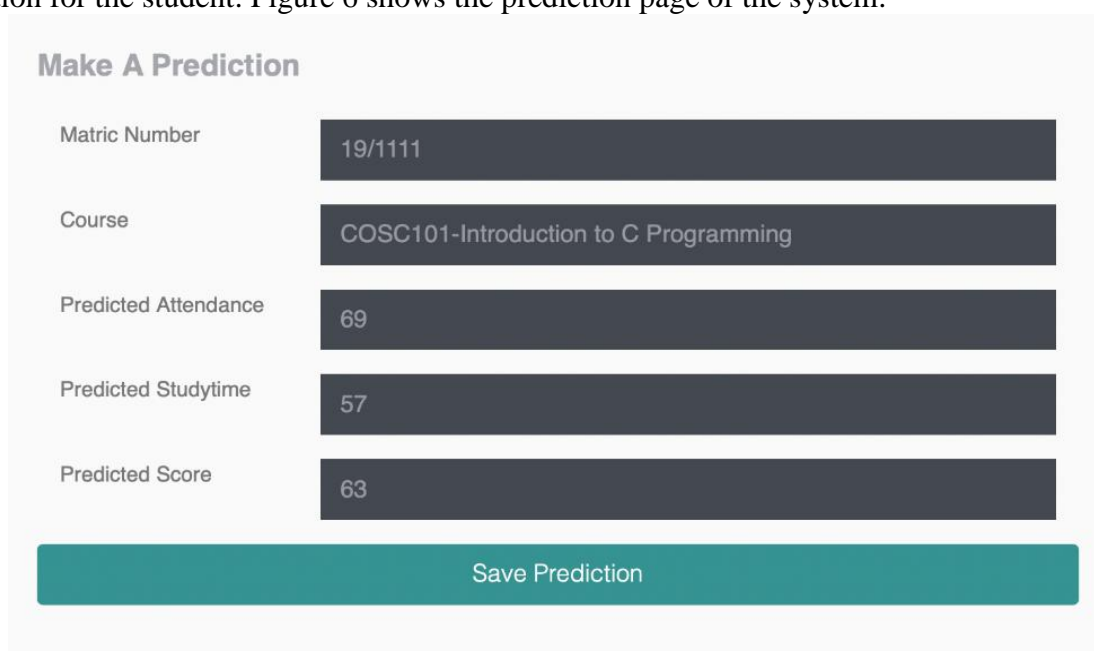

Figure 6. Saving a predicted grade

5. CONCLUSION AND RECOMMENDATIONS

The web-based grade prediction system was validated and verified. It meets up to the goals that were initially set at the onset of the research. The objectives of this study as initially stated have been met. With the use of a web-based grade prediction system which will reduce the stress of the staffs and making it user friendly. The workload of making the exam scores and compiling the semesters work will be reduced. A system that fully automates the entire process of the grading system will without a single doubt make a positive impart in any university.

This application is recommended to all the universities in Nigeria. Based on the challenges encountered during the course of this study, there are areas that can be further worked on to make the system much more available with a wider range of application resources embedded within the system. A situation is expected whereby there will be need for further optimization of the system. As with any online or offline application that requires security measures to prevent and reduce unauthorized intrusion into the application. We appreciate any remarks given towards enhancing the performance of the application and the use in universities across Nigeria. The designed predictive grading system is still open and well-adapted for further modifications.

REFERENCES

[1]   Nithya P, Umamaheswari B, & Umadevi A (2016) A survey on educational data mining in field of education. *Journal of Computer Science & Software Development*; Chapter 1:1-6.

[2]   Sarker, F., Thanassis, T., & Hugh, C.D (2013) Student's performance prediction by using institutional internal and external open data sources. CSEDU: *5th International Conference on Computer Supported Education, Germany.*

[3]   Oyerinde, O. D., & Chia, P. A. (2017). Predicting students' academic performances–
A   learning analytics approach using multiple linear regression. *International Journal of Computer Applications*, 157 (4), 37-44

[4]   Bharadwaj, B.K., & Pal., S. (2011). Mining Educational Data to Analyze Students' Performance, *International Journal of Advance Computer Science and Applications (IJACSA)*, 2(6), 63-69.

[5]   Holmes, G., Donkin, A., & Witten, I.H. (1994). "Weka: A machine learning workbench". *Proc Second Australia and New Zealand Conference on Intelligent Information Systems, Brisbane, Australia.* Retrieved 2021-09-25.

[6]   Kember, D. (1995). Open learning courses for adults: A model of student progress, Englewood Cliffs, NJ: Education Technology.

[7]   Quadri, M.N., & Kalyankar, D. N. V. (2010). Drop Out Feature of Student Data for Academic Performance Using Decision Tree., *Global Journal of Computer Science and Technology*, 10(2), 8-20.

[8]   Quinlan, R. (1993). C4.5: *Programs for Machine Learning*, Morgan Kaufmann Publishers, San Mateo, CA.

[9]   Yadav, S.K., Bharadwaj, B.K., & Pal, S. (2011). Data Mining Applications: A comparative study for predicting students' performance, *international journal of Innovative Technology and Creative Engineering (IJITCE)*, 1(12), 34-43.

[10]   Obsie, E. Y., & Adem, S. A. (2018). Prediction of student academic performance using neural network, linear regression and support vector regression: a case study. *International Journal of Computer Applications*, *180*(40), 39-47.

[11]   Yadev, S.K., & Pal., S. (2012) Data Mining: A Prediction for Performance Improvement of Engineering Students using Classification, *World of Computer Science and Information Technology (WCSIT)*, 2(2), 51-56.

[12]   Ian H. Witten; Eibe Frank; & Mark A. Hall (2011). "Data Mining: Practical machine learning tools and techniques, 3rd Edition". Morgan Kaufmann, San Francisco. p. 191.

[13]   Siti Dianah Abdul Bujang., Ali Selamat., Ondrej Krejcar (2021). IOP Conf. Ser.: Mater. Sci. Eng. 1051 012005

[14]   Hussain, S., Khan, M.Q. (2021) Student-Performulator: Predicting Students' Academic Performance at Secondary and Intermediate Level Using Machine Learning. *Ann. Data. Sci.* https://doi.org/10.1007/s40745-021-00341-0

[15]   Ayşe Yılmaz (2018) Predicting student performance based on summaries and (academic) English skills.

[16]   Yağcı, M. (2022). Educational data mining: prediction of students' academic performance using machine learning algorithms. *Smart Learn. Environ.* 9, 11. https://doi.org/10.1186/s40561-022-00192-z.

[17]   Iqbal, Zafar., Qadir, Junaid., Mian, Adnan., and Kamiran, Faisal. (2017). Machine Learning Based Student Grade Prediction: A Case Study.