Student Performance Prediction In Online Courses Using Machine Learning Algorithms

D J Samatha Naidu[1], K.Venkata Ramya [2], D. Satheesh[3]

Principal of Annamacharya PG of Computer Studies, Assistant Professor of Annamacharya PG College of Computer Studies, Student of Annamacharya PG College of Computer Studies.

Abstract: Advances in Information and Communications Technology (ICT) have increased the growth of Massive open online courses (MOOCs) applied in distance learning environments. Various tools have been utilized to deliver interactive content including pictures, figures, and videos that can motivate the learners to build new cognitive skills. High ranking universities have adopted MOOCs as an efficient dashboard platform where learners from around the world can participate in such courses. The students learning progress is evaluated by using set computermarked assessments. In particular, the computer gives immediate feedback to the student once he or she completes the online assessments. The researchers claim that student assessment

Index Terms - **Massive Open Online Courses (MOOCs), Open University Learning Analytics Dataset (OULAD),Receiver Operating Characteristic(ROC).**

I. INTRODUCTION:

Massive Open Online Courses (MOOCs) is one of the most widespread e-learning platforms. The MOOCs present the course using digital tool materials in various forms such as visual, audio, video and plain text. Most students prefer using video lectures to understand the contents of lessons over thoroughly reading plain text documents. The interactive video in the MOOCs could reduce students' stress, help them to feel relaxed and learn quickly. MOOCs can be classified into two distinct types mainly, connectivist Massive Open Online Courses (cMOOCs) and eXtended Massive Open Online Courses (xMOOCs).The xMOOCs are learning paradigm based on the principles of cognitivist behaviorist theory.

The video lectures featuring the course instructor reviewing the content of the previous online lesson are released weekly. The participants can watch and pause the video at their own pace. Moreover, the students can socially interact with other participants and the instructor through posting in discussion forums. The instructors usually post questions, provide task solutions and reply to student questions via these discussion forums; as a consequence the discussion forums play a vital role in enhancing the course quality and make online sessions collaborative and engaging. The cMOOCs are a new learning model based on connectivist learning theory

2, EXISTING SYSTEM:

The Factor Analysis Model (FAM) was proposed to predict the student's performance in Intelligent Tutoring System (ITS) taking into consideration the difficulty level of assessments based on Item Response Theory concept [9] [10]. The difficulty level of tasks can infer measurement of the correlation between the student's performances and assessment questions. To compute the probability of a student solving a task correctly, a set of predictor variables are defined in the FAM including the number of opportunities presented to the student at each task, the duration spent on each step and the difficulty level of each question or latent variable. The results reveal that incorporating the latent variables into the estimates of student performance can significantly enhance the model [10].

To measure how the activities of learners could impact their learning achievement in MOOCs, the researchers found that Learning Analytics (LAs) in conjunction with machine learning, are effective tools that offer the potential to trace student knowledge. The researchers demonstrated that machine learning could help the educator in providing cohort information about the learning process, furnishing researchers with the ability to both visualise and analyse the information obtained from each tier of the learner. Thus, an accurate predictive model can be acquired in such courses[11] [12][13]. Students' marks in the first assessment and quiz scores in conjunction with social factors are used to predict students' final performance in online course [14].

Disadvantages:

- In the existing work, the system is poor performance in which the assessments are not Tutor Marked Assessment (TMA), Computer Marked Assessment (CMA).
- This system is less performance due to Lack of Massive open online courses (MOOCs)..

PROPOSED SYSTEM:

The OULAD dataset was captured from the Open University Learning Analytics Dataset (OULAD)repository. The open university in the UK delivers the online course in various topic for undergraduate and postgraduate students in the period between 2013-2014. The main composite table called "studentInfo" is linked to all tables. The"studentInfo "table includes information relevant to students' demographic characteristics. The information related to students performance are collected in "Assessments" and Student Assessment tables. The table "Assessments" contains information about the number, weight and the type of assessments required for each module. In general, each module involves a set of assessments, followed by the final exam. The assessments are Tutor Marked Assessment (TMA), Computer Marked Assessment (CMA). The final average grade is computed with the sum of all assessments (50%) and final exams (50%). The "Student Assessment" table involves information relating to student and the assessment mark. The "Student Registration" table contains information about the date the students registered and unregistered in a particular module. The overall date is measured by counting numbers of unique days that students interact with courses until the course ends. In Open University online courses, students are able to access a module even before being a student of the course; however, it is not possible to access the course postcourse closure date.

Advantages:

- Students' marks in the first assessment and quiz scores in conjunction with social factors are used to predict students' final performance in online course.
- Learning Analytics (LAs) in conjunction with machine learning, are effective tools that offer the potential to trace student knowledge.

RESEARCH METHODOLOGY:

The OULAD dataset was captured from the Open University Learning Analytics Dataset (OULAD)repository. The open university in the UK delivers the online course in various topic for undergraduate and postgraduate students in the period between 2013-2014. The main composite table called "studentInfo" is linked to all tables.The"studentInfo "table includes information relevant to students' demographic characteristics[15]. The information related to students perfromnce are collected in "Assessments" and Student Assessment tables. The table "Assessments" contains information about the number, weight and the type of assessments required for each module. In general, each module involves a set of assessments, followed by the final exam. The assessments are Tutor Marked Assessment (TMA), Computer Marked Assessment (CMA). The final average grade is computed with the sum of all assessments (50%) and final exams (50%). The "Student Assessment" table involves information relating to student 2022 International Joint Conference on Neural Networks (IJCNN),Issue Date:28.September.2022 assessment results, such as the date of the submitted assessment and the assessment mark[15]. The.

Students' performance Model

Two sets of experiments are conducted in this case study. In the first experiment, the dynamic behavioral features are considered to predict student performance, while the static behavioral attributes are employed in the second experiment.

The problems are formulated as classification and regression. The regression setting is considered when we aim to predict students' assessments grades, whereas classification setting is utilised when we seek to predict final student performance in the entire course. It is considered a multi-class problem where the target class is whether students pass, fail or withdraw from courses.
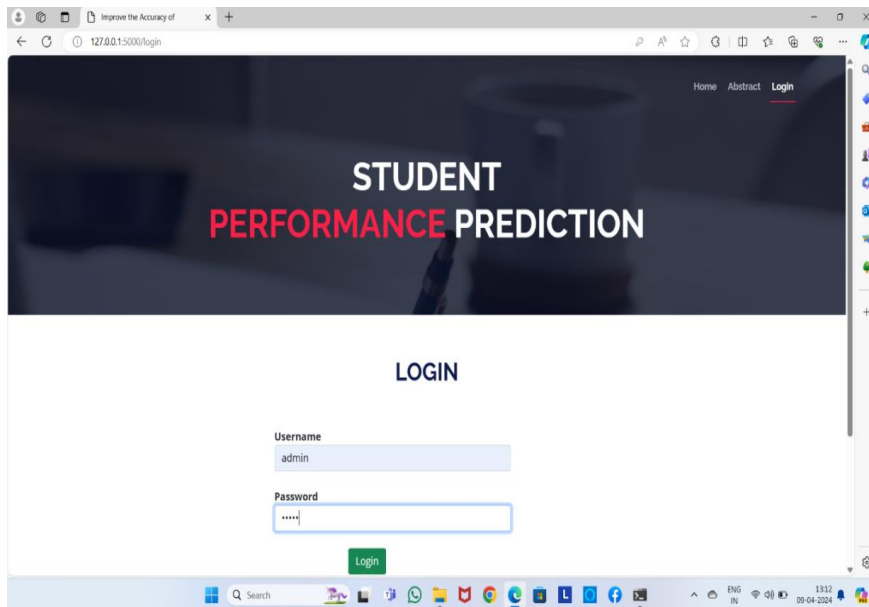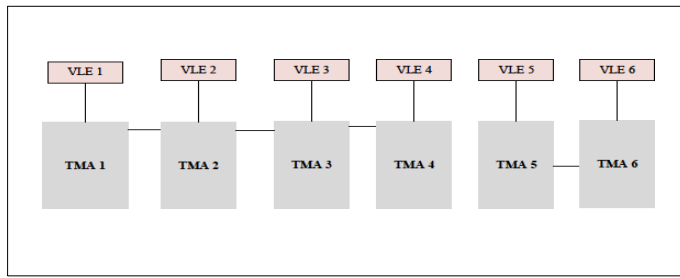
.

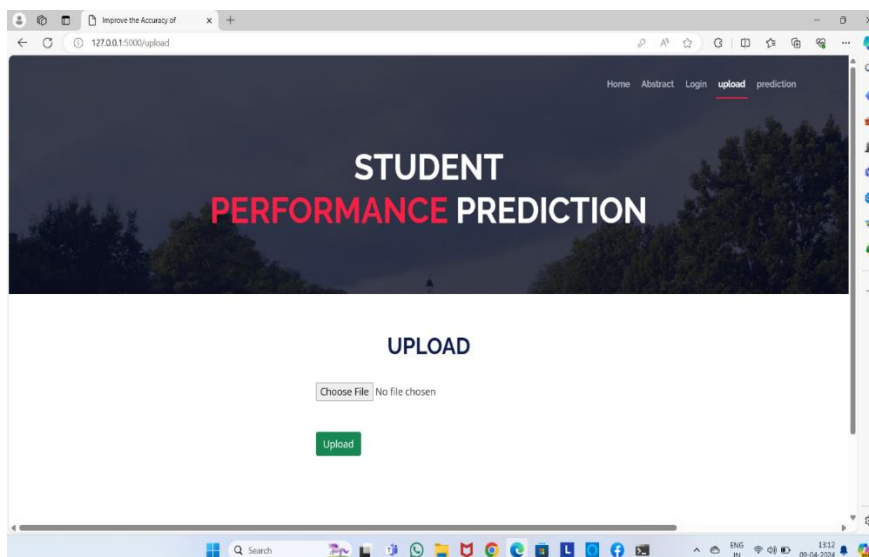**fig:** it is a login interface through this we can login in to the software

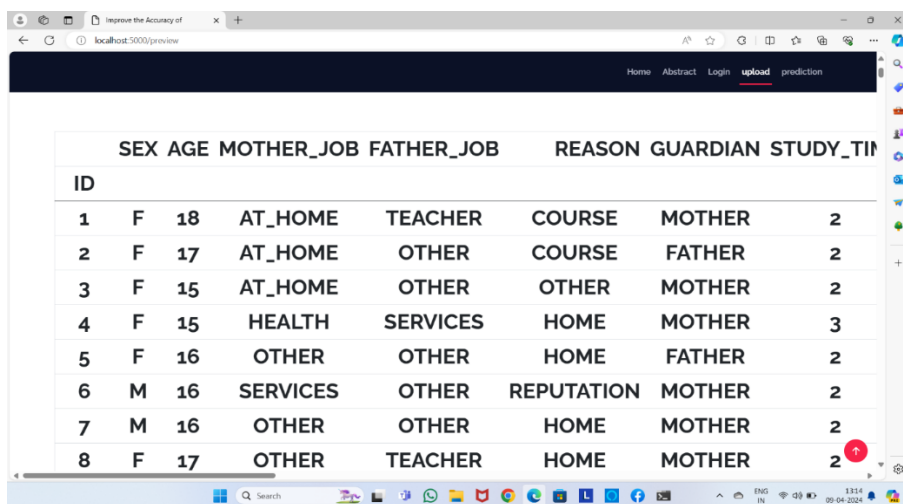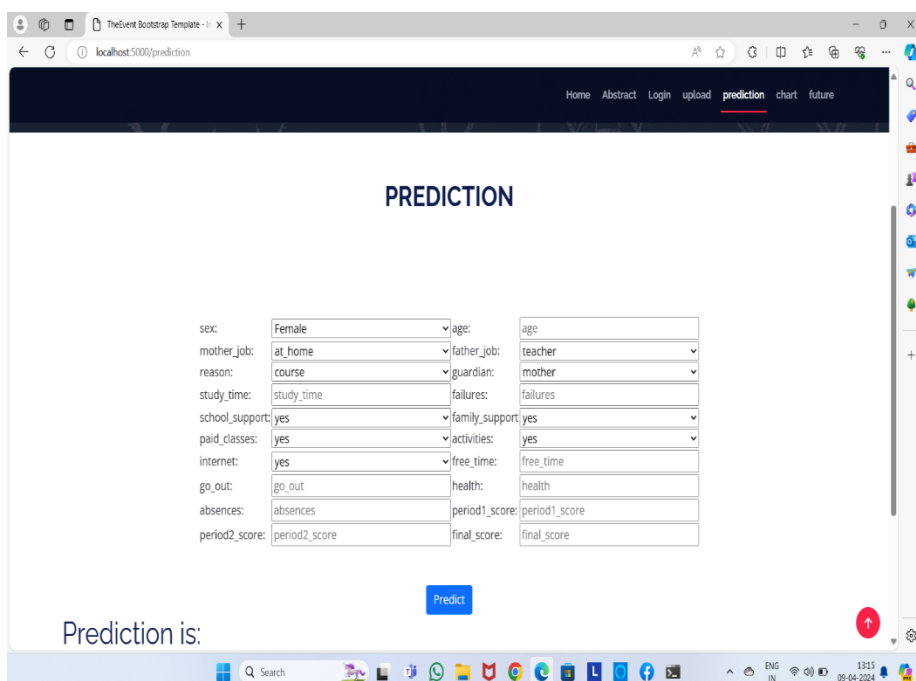**Fig:** here in this we have to upload the training data for the prediction process

Fig: this is the upload data collected from various source

Fig: this is the student performance prediction interface we have to fill the table to predict student performance

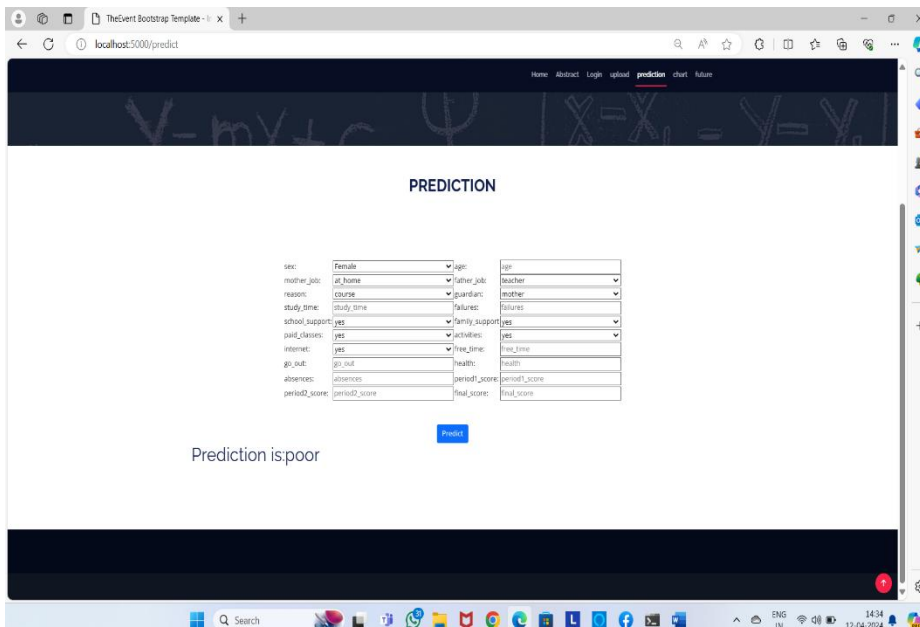**fig:** this is an example output of a student that he got good in the prediction

fig:this is an example output of a student that he got poor in the prediction
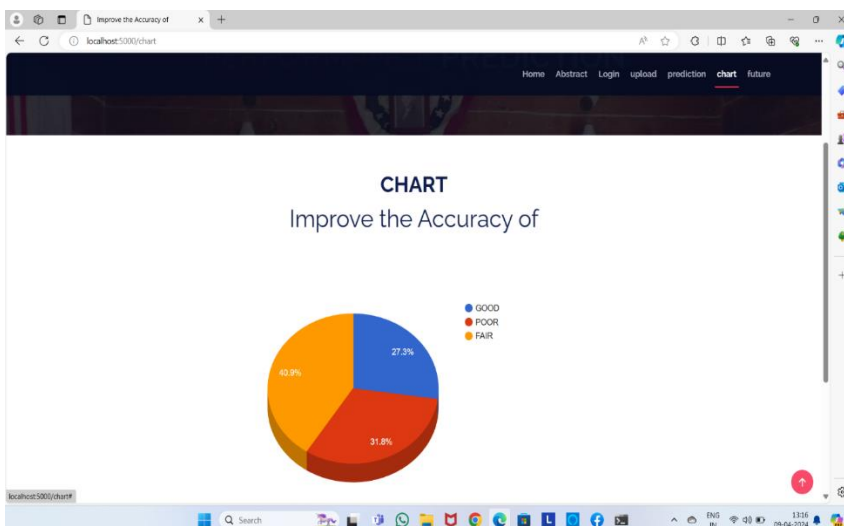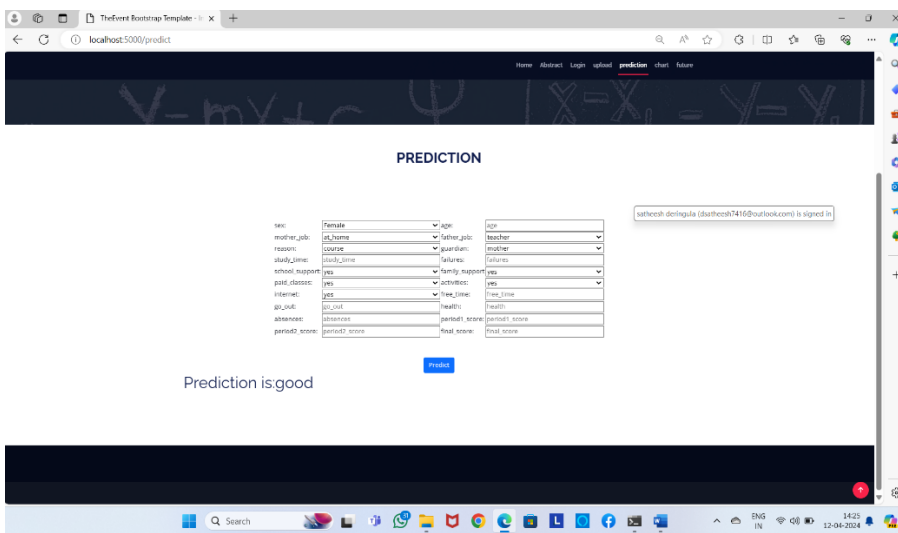

Fig: this is the prediction pie chat that describe the percentage of the result

Acknowledgment:

Two sets of exterminates have been carried out in this study using regression and classification analysis. The results of predicting students' assessments grades model show that the students' performance in a particular assignment relies on students' mark in the previous assignment within single Courses. The researchers conclude that students' prior grade point average (GPA) with a low mark is considered as a significant factor of withdrawal from the next course in the traditional classroom setting,

REFERENCES:

1. K. F. Hew and W. S. Cheung, "Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges," *Educ. Res. Rev.*, vol. 12, pp. 45–58, 2014.

2. H. B. Shapiro, C. H. Lee, N. E. Wyman Roth, K. Li,

3. M. Çetinkaya-Rundel, and D. A. Canelas, "Understanding the massive open online course (MOOC) student experience: An examination of attitudes, motivations, and barriers," *Comput. Educ.*, vol. 110, pp. 35–50, 2017.

4. J. Renz, F. Schwerer, and C. Meinel, "openSAP: Evaluating xMOOC Usage and Challenges for Scalable and Open Enterprise Education.," *Int. J. Adv. Corp. Learn.*, vol. 9, no. 2, pp. 34–39, 2016.

5. S. Li, Q. Tang, and Y. Zhang, "A Case Study on Learning Difficulties and Corresponding Supports for Learning in cMOOCs| Une étude de cas sur les difficultés d'apprentissage et le soutien correspondant pour l'apprentissage dans les cMOOC," *Can. J. Learn. Technol. Rev. Can. l'apprentissage la Technol.*, vol. 42, no. 2, 2016.