ISSN: 2322-3537
Vol-13 Issue-01 Feb 2024

# Predicting Students' Success in Online Courses with the Use of Machine Learning Algorithms

**V.Chiranjeevi, Dr. K.Spurthy, Syed Parveen, N.Kamala Vikasini**

[2] ASSOC.PROFESSOR, [1,3,4] ASSIT.PROFESSOR

Department of CSE, SWARNA BHARATHI INSTITUTE OF SCIENCE & TECHNOLOGY (SBIT), Pakabanda Street, Khammam - 507 002. Telangana, India.

## ABSTRACT

Developments in Expression and Specification Modern developments in information and communication technology have led to the proliferation and widespread use of massive open online courses (MOOCs) in the field of online education.Various methods have been used to provide interactive material, such as pictures, numbers, and movies, with the goal of motivating pupils to cultivate new cognitive abilities. Massive open online courses (MOOCs) are being offered by some of the most famous institutions in the world to students all over the world. Predetermined, computer-marked assessments are used to evaluate students' progress. Specifically, the computer instantly gives feedback as soon as the student finishes an online test. The authors of the research claim that the likelihood of students succeeding in an online course is related to their involvement and performance in the previous session. The subject of whether or not students' prior test performance and involvement could influence their future exam performance has received little attention in the literature. Based on the results of the evaluation and the efficiency of the final trainees, two anticipatory versions have been created in this study. These designs will allow us to identify what aspects of MOOCs are most important for students to succeed in when it comes to discovery learning. As can be seen from the results, both models provide quite practical results. While GBM yielded the most accurate findings for the students' final performance, with an average value of 0.086, the most cost-effective RSME gain for the students' analysis grades model was 8.131 for RF.

***Keywords: MOOC, PIC, ML, DL, RSME, GBM, RF.***

## INTRODUCTION

The Massive Open Online Courses (MOOCs) are among the most popular kinds of online education. The training courses offered by massive open online courses (MOOCs) make use of a variety of electronic tool items, including visual, audio, video, and plain text. Rather of reading lengthy plain-text publications, many students find that video clip speeches help them better comprehend course content. Interactive videos in massive open online courses (MOOCs) have the potential to ease students' minds, make them feel more at ease, and speed up their learning. the first two There are two main categories of massive open online courses (MOOCs): eXtended large open online courses (xMOOCs) and connectivity massive open online courses (cMOOCs). The principles of cognitive behaviourism are the foundation of the new paradigm that the xMOOCs are revealing. [4] The programme are structured similarly to the typical training course, with a final exam, a series of multiple-choice quizzes, and video lectures making up the curriculum. Once a week, students may see video lectures in which the training course instructor goes over the material from the previous online session. People may watch the film at their own leisure and pause it whenever they choose. In addition, by posting in discussion boards, students may socialise with the instructor and other participants. Discussion online forums play an important part in improving the course quality and making online sessions collaborative and attractive since instructors generally express problems, suggest task solutions, and react to student complaints via these forums. the third [5] A new iteration of massive open online courses (MOOCs) based on connectives theory of learning [3] [4] Trainees acquire the training course curriculum by asking questions and sharing this information with other participants using the collectivism approach; the instructor does not provide the actual knowledge material. Works Cited [3] [4] Assume for the sake of argument that massive open online courses (MOOCs) rely on a collaborative approach to discovery, with students working together to create a final output that is both reusable and remixable before sending it on to future students. Unlike xMOOCs, where college lecturers may utilise computer-marked analytical replies to gauge students' knowledge, cMOOCs make it hard to include expertise in assessing students' comprehension. In particular, when the student completes the online analysis, the computer system provides immediate feedback. Upon successful completion, the student will undoubtedly get their xMOOC accreditation.There is no official assessment in the cMOOCs. This is why massive open online courses (cMOOCs) are not an option for academic institutions. [5] [6]

The rapid development of AI in recent years has made it a viable option for screening and evaluating students' performance in online courses. There have been few efforts to evaluate the performance of the trajectories, despite the fact that some researchers have utilised machine learning to predict student success in [7]. [8] Consequently, educators are unable to monitor students' actual learning curves in real time. This paper conducts two sets of experiments. Students' assessment ratings are estimated using regression analysis in the first set of experiments. To forecast the trainee's outcome, we look at their present and previous actions, as well as their performance. The second round of trials aims to predict students' long-term performance using a monitored machine learning strategy. There are three types of prospect predictors that have been considered: behaviour functions, group traits, and temporal functions. The proposed revisions let teachers keep tabs on students' immediate performance and give fresh understanding of the most important learning activity. To the best of our knowledge, the online programme has only ever used two outcomes—"success" and "fail"—to evaluate student achievement. "Success," "fail," and "took out" are the three-class labels that our version uses to predict the performance.

Based on the notion of Product Action Theory, the Element Evaluation Design (FAM) was suggested as a way to forecast the trainee's performance in an Intelligent Tutoring System (ITS) by considering the degree of difficulty of evaluations. [9] [10] One facet of the relationship between students' performance and evaluation queries may be inferred from the difficulty level of assignments. The FAM specifies a collection of forecaster variables that include the amount of chances given to the student for each assignment, the time spent on each action, and the difficulty level of each inquiry or hidden variable in order to calculate the possibility of a student correctly completing a task. The findings show that the version may be much improved by including the hidden factors into the student performance estimations. [10]

## SURVEY OF RESEARCH

Researchers have shown that a combination of Discovering Analytics (LAs) and machine learning provides a promising way to map trainee understanding, which might help them assess the impact of learner activities on their understanding success in massive open online courses (MOOCs). Scientists were able to conceptualise and assess data collected at each level of the student's understanding process, proving that AI may aid the teacher in providing friend details on the learning process. As a result, these programmes can provide an accurate rendition of the forecast. the eleventh [12] in [13] Students' social factors, exam scores, and first-analysis grades are used to forecast how well they would do in an online class. [14]

We introduced two models that can predict outcomes. Whether trainees obtained a standard or distinction certification was predicted in the first edition using logistic regression. The second design also used logistic regression to predict whether students met the qualifying criteria or not. The results showed that the amount of peer reviews is a significant factor in determining the quality of the results. In order to get a certificate, average test results were thought to be a reliable indicator. With a percentage of 92.6% for the first model and 79.6% for the second design, particularly, the accuracy of differentiation and regular versions were recorded. [14]

Researchers at UMBC looked at the correlation between student performance and data collected from virtual learning environments (VLEs) [12]. It was LA was achieved by use of the Inspect My Activity (CMA) application. CMA is a LA tool that provides regular feedback on trainees' feelings and compares their VLE activities to others. Students who participated in the programme more often compared to those who did not to have a higher likelihood of receiving a grade of C or above [12].

## PROPOSED SYSTEM

### Synopsis of Contents

The OULAD dataset was extracted from the OULAD database, which stands for Open University Understanding Analytics Dataset. Online training courses in a variety of subjects are available to undergraduate and graduate students from the Open University in the United Kingdom during the academic year 2013–2014. "Trainee Info" is the primary composite table that is linked to all the tables. The "student Info" table includes data that is pertinent to the demographic characteristics of students. [15] Pupil Analysis and "Evaluations" tables compile data pertaining to students' academic progress. For each module, you may find the necessary number, weight, and kind of assessments in the "Assessments" table. In most cases, the final exam follows a set of analyses that are included in each component. Both Tutor Marked Analysis (TMA) and Computer System Marked Analysis (CMA) are part of the evaluations. All tests (50%) and final exams (50%) are used to calculate the last ordinary quality. Trainee and analysis mark data is included in the "Student Analysis" table. [15] The "Trainee Registration" table contains information on the trainees' sign-up dates, regardless of whether they are included in a specific module or not. The total number of days is determined by keeping track of the number of separate days that students access the programmed until the training course concludes. While students enrolled in an Open University online course may view a module before the start of the training, they will no longer have access to the programme after the course has ended. Data collected from students' Virtual Knowing Atmospheres pertains to their interactions with technological devices.

This case study includes the results of a number of experiments designed to improve students' efficiency using the Model 2. The first experiment uses characteristics of students' dynamic behaviour to predict how well they would do in class, whereas the second uses characteristics of students' static behaviour. The issues manifest as regression and categorization. When aiming to forecast students' analytical abilities, the regression setup is considered; when aiming to forecast students' overall programme performance, the classification setup is used. Considered here is a multi-class problem whose goal is the success or failure of training courses. Instructors might be able to help students with poor ratings more quickly by using early quality prediction to find further discoveries goods and treatment support. [7] The learner is required to complete the final exam in addition to the five CMA analyses and six TMA assessments that have already been covered. There is a strict deadline for submitting the analyses. Our temporal analysis is based upon the TMA entrance date since the TMA assessment considers 45% of the final outcome while the CMA analysis only considers 5%. Students' efficiency is quickly asserted in the initial set of trials. As a result, the training course is divided into six time periods, with evaluation entry days relating to each. On the day of the assessment, the students' behaviour records are handed out. This evaluation takes into account the student's performance on previous assessments related to their communication acts.

We combined the child's behavioural tasks from all six time pieces into one single time piece to evaluate the trajectories of pupil efficiency in the second set of studies. As input variables, we make use of the behavioural, market, and temporal functions. Our calculations did not take into consideration the grades received on previous exams or the final test score, which are factors in determining the target course. Out of 4004 papers in the dataset, 28% fall into the "stop working" category, 40% into the "withdrawn" category, and 32% into the "pass" category.

**The selection of features**

We will only analyse the included option during the first set of experiments since our aim is to explore how child performance in the previous assessment affects the following evaluation. This case study makes use of Recursive Function Elimination (RFE). One of the most popular methods for selecting wrapper functions is reclusive feature elimination. You might think of RFE as an optimisation algorithm that uses resampling and reverse selection as its foundation. Up until it obtains a minimal set of characteristics, it continues to create the model iterative.

International Journal of applied exercise physiology
ISSN: 2322-3537
Vol-13 Issue-01 Feb 2024

There are a number of components that distinguish the data set into train and bootstrap samples. We choose the algorithms at each model based on their importance. We keep using the new design, which includes one of the most important forecasts, to test the probability of ranking functions until we're all exhausted.

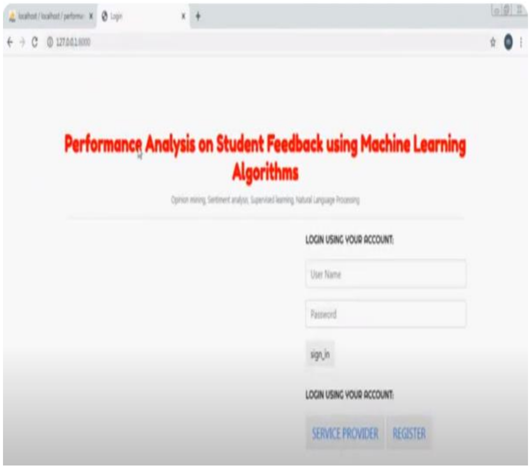


Fig.1. Home page.

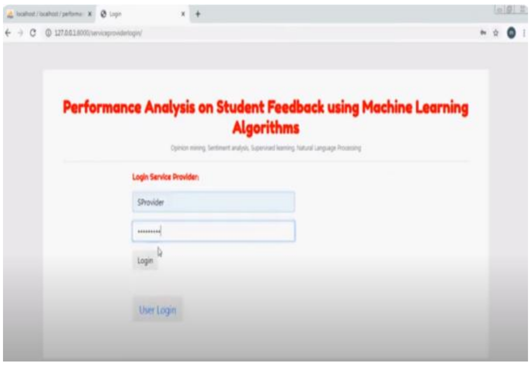


Fig.2. Admin login page.

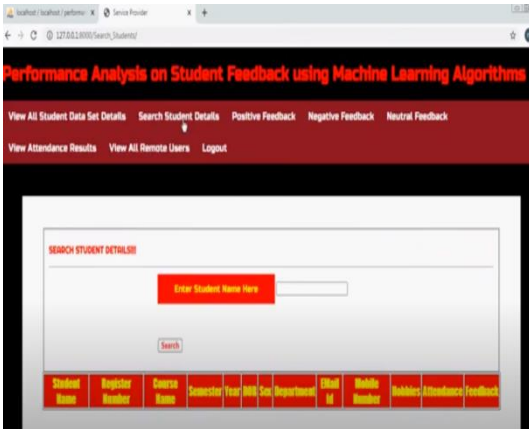


Fig.3. Search student details.

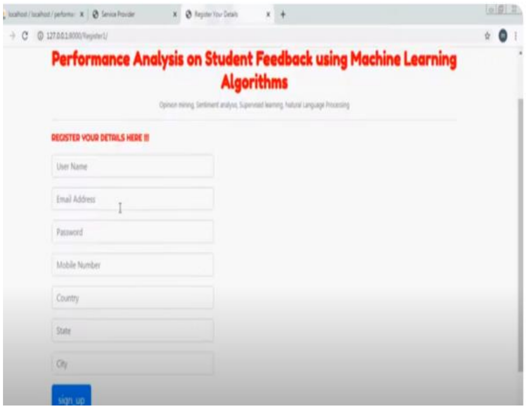


Fig.4. Registration details.

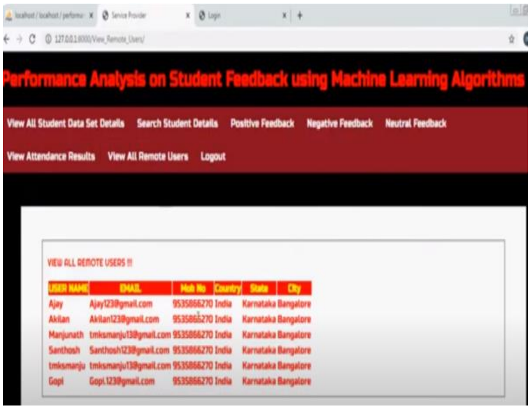


Fig.5. Users details

ISSN: 2322-3537
Vol-13 Issue-01 Feb 2024

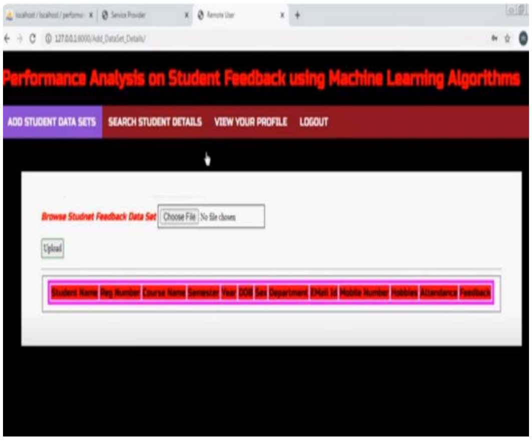


**Fig.6. upload the data set.**

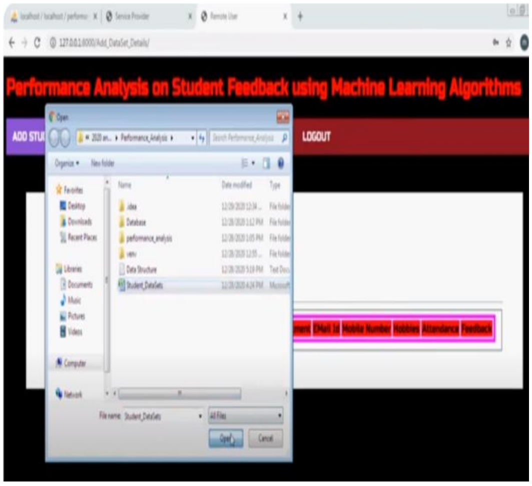


**Fig.7. Load dataset.**

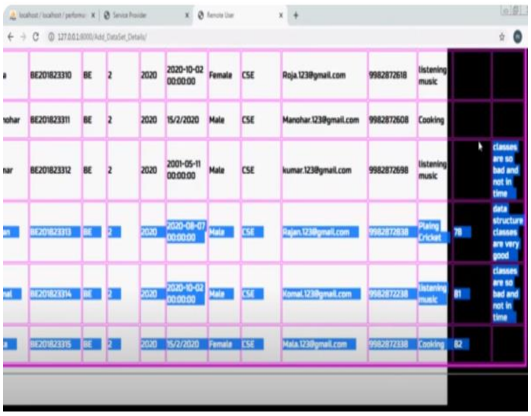

**Fig.8. Student details.**

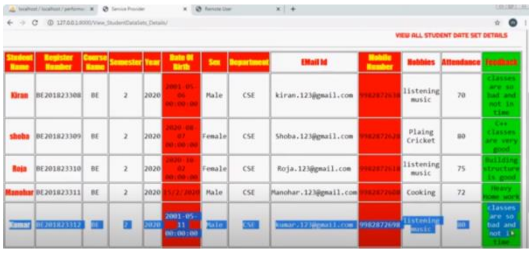

**Fig.9. Output results.**

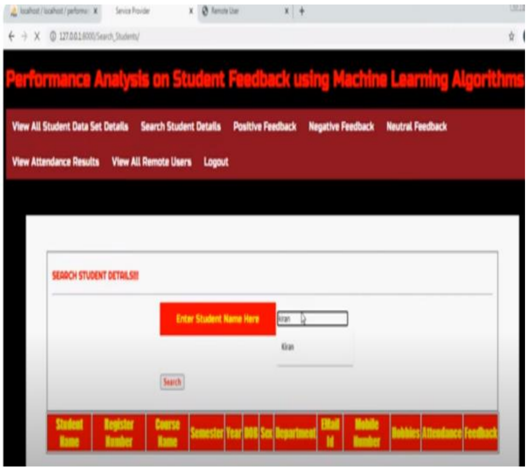


**Fig.10 Output with student details.**

## CONCLUSION

In this research, two sets of analyses were conducted using regression and category assessment. The results of designing assessments based on students' attributes show that, in solo programme, students' performance on one assignment is dependent on their performance on another. The study's authors draw the conclusion that, in a traditional classroom setting, students are more likely to drop out of subsequent classes if their prior grade point average (GPA) is poor. In terms of the reliability of past performance into

ISSN: 2322-3537
Vol-13 Issue-01 Feb 2024

future understanding achievement, traditional and online classrooms are similar.

Students' engagement with course materials is shown to have a major role in their overall achievement, according to the most recent predictive version of student performance. Also, since temporal functions aren't a part of regression assessment, the results demonstrate that long-term trainee efficiency achieves far superior accuracy than students' evaluations grading prediction design. A useful predictor that is highly associated with trainee efficiency is the day of student registration from the training course. The data does not include the latest date of students' activity before taking the assessments, which is a problem with the regression evaluation. Consideration of time elements on expecting of subsequent assessment grades has actually been suggested for the searches. One potential avenue for future research is to utilise temporal features to predict how students will do on examinations. Potentially, more sophisticated machine learning will be used in place of time collection assessment when dealing with temporal attributes.

**ACKNOWLEDGMENT**

We thank CMR Technical Campus for supporting this paper titled "**STUDENTS PERFORMANCE PREDICTION IN ONLINE COURSES USING MACHINE LEARNING ALGORITHMS**", which provided good facilities and support to accomplish our work. I sincerely thank our Chairman, Director, Deans, Head Of the Department, Department Of Computer Science and Engineering, Guide and Teaching and Non- Teaching faculty members for giving valuable suggestions and guidance in every aspect of our work.

## 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, 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.

[3] 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.

[4] 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.

[5] S. Zutshi, S. O'Hare, and A. Rodafinos, "Experiences in MOOCs: The Perspective of Students," Am. J. Distance Educ., vol. 27, no. 4, pp. 218–227, 2013.

[6] Z. Wang, T. Anderson, L. Chen, and E. Barbera, "Interaction pattern analysis in cMOOCs based on the connectivist interaction and engagement framework," Br. J. Educ. Technol., vol. 48, no. 2, pp. 683–699, 2017.

[7] W. Xing and D. Du, "Dropout Prediction in MOOCs: Using Deep Learning for Personalized Intervention," J. Educ. Comput. Res. p.0735633118757015., 2018.

[8] M. J. Gallego Arrufat, V. Gamiz Sanchez, and E. Gutierrez Santiuste, "Trends in Assessment in Massive Open Online Courses," Educ. Xx1, vol. 18, no. 2, pp. 77–96, 2015.

[9] B. J. Hao Cen, Kenneth Koedinger and Carnegie, "Learning Factors Analysis – A General Method for Cognitive Model Evaluation and Improvement," in In International Conference on Intelligent Tutoring Systems, 2006, vol. 8, pp. 164–175.

[10] K. R. Koedinger, E. A. Mclaughlin, and J. Stamper, "Automated Student Model Improvement," in Educational Data Mining, proceedings of the 5th International Conference on, 2012, pp. 17–24.

[11] J. Sinclair and S. Kalvala, "Student engagement in massive open online courses," Int. J. Learn. Technol., vol. 11, no. 3, pp. 218–237, 2016.

[12] J. Mullan, "Learning Analytics in Higher Education," London, 2016.

[13] P. and K. Al-Shabandar, R., Hussain, A.J., Liatsis, "Detecting At-Risk Students With Early Interventions Using Machine Learning Techniques," IEEE Access, vol. 7, pp. 149464–149478, 2019.

ISSN: 2322-3537
Vol-13 Issue-01 Feb 2024

[14] S. Jiang, A. E. Williams, K. Schenke, M. Warschauer, and D. O. Dowd, “Predicting MOOC Performance with Week 1 Behavior,” in Proceedings of the 7th International Conference on Educational Data Mining (EDM), 2014, pp. 273–275.

[15] L. Analytics and C. Exchange, “OU Analyse : Analysing at - risk students at The Open University,” in in Conference, 5th International Learning Analytics and Knowledge (LAK) (ed.), 2015, no. October 2014.