Assessment Disparities in MOOCs: A Data-Driven Study

Harshita Mogha   
*Amity University Uttar Pradesh*   
harshitamogha2001@gmail.com

Nitasha Hasteer

*Amity University Uttar Pradesh*  
nitasha78@gmail.com Hasan Kamal  
*Amity University Uttar Pradesh* hasankml87@gmail.com

K M Soni

*Amity University Uttar Pradesh*

kmsoni@yahoo.comIsra Imtiyaz   
*Amity University Uttar Pradesh*

israimtiyaz9@gmail.com

*Abstract*— Massive Open Online Courses (MOOCs), are online courses that are accessible and adaptable and are intended for massive enrolments. They offer low-cost access to high-quality education, self-paced learning and a variety of subject options that can improve job prospects. Through platforms like Coursera, edX, and NPTEL, universities and other organisations provide MOOCs, enabling learners all over the world to access high-quality education. Without the limitations of conventional classroom instruction, MOOCs also assist professionals in improving their skills and staying current with industry developments. Therefore, it is crucial to understand the difference in the performance of technical and non-technical students in different MOOCs to effectively design pedagogical tactics in today's rapidly changing educational landscape. This study is aimed at investigating the performance of students on internal and external assessments from two different backgrounds i.e. technical and non-technical MOOCs. K-means clustering is used to compare the results of students from different courses. The result comprises of internal and proctored test scores of two different courses in a dataset. Clustering was performed with careful data preparation, including outlier removal to get the accurate results. We demonstrate how students from different backgrounds behaved in analogous assessment situations by contrasting the clustered groups. Our results indicate different performance patterns between technical and non-technical students with substantial implications on higher education's curriculum development and resources for supporting student needs. The variation in performance emphasizes a need for varied teaching designs tailored to different learning approaches than present in non-technical classes.

# Introduction

MOOCs have transformed education by enabling everyone with an internet connection to access top-notch learning materials. This provision is led by some of the world's best colleges and institutes and includes a vast range that would run from computer science to humanities. This has contributed to students' massive intake in such courses, basically increasing the accessibility and flexibility of education.

Evidently, MOOCs have influenced the present landscape in education. By eliminating regional restrictions, they have democratized education and made learning more accessible to everyone at a lower cost. Students can start their journey, review lectures, or join a worldwide community of learners at their own pace. Flexibility was particularly advantageous for working professionals, stay-at-home parents, and individuals living in distant places who cannot access traditional educational institutions[1].

MOOCs have also become more assimilated in their formal education frameworks. Many universities and schools have been aware of the potential brought about by MOOCs and are incorporating them into their teachings. Of late, some colleges even grant credit in completing MOOCs as an alternative to traditional classroom instructions[2]. As more organizations and educators acknowledge MOOCs as useful educational resources that assist students in establishing themselves in the contemporary world, this trend is likely here to stay [3].

MOOCs stand as one of the greatest educational innovations in recent years. Because they address several issues and offer opportunities, MOOCs have recently gained a lot of attention in the educational community. With MOOCs increasingly being implemented in schools and colleges, their utility and impact on student achievement need to be ascertained. Access to quality education from anywhere in the world can change the way we think about learning and teaching.

With a number of factors propelling this expanding MOOCs trend, student enrolment in these courses has skyrocketed, making education more flexible and accessible. MOOCs have enhanced accessibility and inclusivity by removing barriers to high-quality education, making it accessible to people all over the world, regardless of location or financial position. This has balanced the playing field in education, allowing students from all backgrounds to register in courses offered by premier colleges and institutes[1]. Because learners could learn at their own speed and schedule, it became extremely attractive to those with no access to traditional system of education. Similarly, flexibility and convenience are characteristics provided by MOOCs. One of the factors contributing to why MOOCs gained attention was its flexibility. Here, one learns independently at his or her pace. Lectures can be watched more than once, projects finished individually, and so forth. Another advantage of MOOCs is that they cost much less than traditional learning. Many MOOCs can be joined for free, with the option to purchase the course for getting certification. This has made education available to a higher proportion of learners, specifically to low-income families that would otherwise can’t afford to join universities[3]. MOOCs provide students with a global network and context through which they can communicate and interact with each other across the globe. Such exposure fosters a feeling of belonging together with collaborative learning. Interaction with experts and professionals in the world is most likely to increase new job development and other professional advancements[4].Continuous learning and skill development are the keys to success in this changing job economy. MOOCs offer a wide range of courses for every new necessity that comes about. From understanding a new programming language to data analysis or soft skill training, MOOCs are there to provide a handy approach to staying updated and competitive on the job market[5].

What presents a significant importance of MOOCs in modern time is their potential to address prominent issues in education while accessible, flexible, and high-quality learning opportunities are given. As technology evolves as well as the need to continue learning, we perceive that MOOCs have much more to do to support the future of learning. Several global universities have already integrated MOOCs into their curriculums.

Since MOOCs are increasingly used in schools and colleges, it is imperative to determine if they are useful and affect the student's achievements. There are many reasons why colleges believe MOOCs can help them cultivate their students' futures. MOOCs are produced by top universities and institutes. As a result, the content produced is mostly of good quality. Such reputations attract students looking for authentic and reliable sources of learning material. Most MOOCs contain quizzes, forums, and peer-reviewed assignments. All these features make learning much more interesting and interactive. This results in better retention and comprehension of the subject[6].

As more and more universities and employers are recognizing the utility of MOOCs, certification and credits acquired through MOOCs are gaining value and thus it makes the student feel that his or her time and efforts used to complete MOOCs will be rewarded. MOOCs balance the academic load so that students can handle their academics much better. The fact that they can study at their own pace, and review lectures and materials, would reduce the possibility of frustration. Thus, it would make the approach to MOOCs appealing and useful[6].

This study aims to undertake statistical analysis of student performances in internal and external evaluations for two different MOOCs. Although MOOCs have been widely noted to facilitate access to higher education, most of the recent research on MOOCs centers on aggregate metrics of performance or specific indicators like participation rates and the final test results. However, there is an inadequate appreciation of how a student might perform in a continuous stream of internal assessment as against an external examination. In this regard, our study examines the relationship between internal assessments and proctored exams to determine whether strong performance in continuous evaluations translates to success in final exams, or if discrepancies indicate the need for different teaching methods or additional support.

The other part of the paper is sectionalised as follows. Section 2: Literature Review, in which we examined influential literature by notable authors relevant to our topic and the methodology is detailed in Section 3.

# Literature Review

In this review, a systematic approach was used to identify and include relevant articles by use of the PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analysis)[7]. From numerous databases, 1,026 entries were identified but after deleting duplicates and applying ineligible records through automation method, only 191 were tested. 98 of the records were excluded based on relevance to the research to be conducted. From the remaining 93 reports whose retrieval was requested, 17 were not acquired, and 20 more were rejected because they were deemed out of scope and involved language barriers. The final tally included incorporation of 56 additional studies, which makes the total count 20 studies to be reviewed. This comprehensive procedure ensured that only the most relevant and high-quality studies were selected, thereby building a firm foundation for comparison between technical students and non-technical students in courses under study.

A diagram of a flowchart

Description automatically generated

Figure 1 PRISMA Diagram

Sahar Voghoei et al. discovered that there is a relationship between the engagement of students in online discussion forums and their performance in university online courses. The main aim of the study was to identify how many students participated in various online course discussion boards during the summer semester of 2019. It tried to determine whether there were any relationships between these rates of engagement and the grades or other academic performances of the students. The study monitored participation throughout the semester and observed for patterns, with specific focus on the timing and the consistency of activity. According to the study, those students whose GPAs are within the 70-80 percentiles were observed with the highest participation throughout the first two-thirds of the semester. These students body were the most active on the discussion forums during the period. The pattern of participation shifted in the last weeks of the semester. It was the students with GPAs between the 87th and 93rd percentiles who participated most toward the end of the third part of the semester. It was determined that it was not necessarily the students who posted the most that were the most successful, but rather those that consistently contributed throughout the course of the semester. One of the major findings was that the top-performing students (those in the 87-93 percentile) did not necessarily post the most overall, but rather regularly participated throughout the semester. Consistent participation, regardless of the number of posts, was recognised as a critical element distinguishing higher-performing students from others. The study did confirm a connection between participation rates and grades; however, it is not a direct correlation where any more than a certain amount of participation equates to the grade[8].

Tianping Deng et al. researched the connection between students' online learning practices in MOOCs and their academic performance. This study contributes to providing insights that can help teachers adapt to teaching practice and better manage the process of supporting students' experiences of learning. Online learning behaviours of students in MOOCs affecting their offline academic performance were measured through final test scores. In an endeavour to investigate these connections, this study collaborated data from MOOCs with scores from regular classroom tests. The experiment draws data from two MOOCs conducted by Huazhong University of Science and Technology: one is for the 2018 Class of Exemplary Engineer Education and the other is for the Advanced Class in Mathematics and Physics. Data included both online and offline activity-for example, video views, test scores, online exam scores, attendance, classroom practices, and number of discussions. The study found that significant learning characteristics, such as classroom practices, number of discussions and number of video views, were highly correlated with final exam scores. Classroom Practices, for instance, have been established to be highly related to the exam performance since they duplicate the exam environment. The number of discussions mirrors actual classroom participation and is positively associated with better academic outcomes. The study revealed that even though attendance is generally promoted, it does not necessarily align strongly with improved academic performance because children can be present in class without contributing actively[9].

Another study was based on an English MOOC course offered by the xuetangx platform, one of the MOOC providers in China. The course was totally online and focused on daily English expressions. It consisted of eight units with 9-10 learning videos and after-school homework and discussion questions. Using the two-stage clustering method, the study confirmed four distinctive patterns of involvement among Chinese MOOC students. Four patterns were categorized into goal-oriented high-engagement, performance-oriented high-engagement, performance-oriented low-engagement, and goal-oriented low-engagement[10].

Another study published in literature explores the link between learner satisfaction with MOOCs and the most prominent reasons that lead to dropout rates. The researchers obtained data from MOOC forums and social media groups that are relevant to the subject. The data was composed of forum postings, comments, and tweets on a Coursera course. Using Support Vector Machines (SVM), the posts made on the forum by learners and on their social media were classified into three categories: satisfied, neutral, and dissatisfied. The mood analysis helped in very clear understanding of the learners' experiences. Most learners stated that they were satisfied or dissatisfied with the MOOCs, whereas very few were neutral. This was essential to understand which aspects were required to be enhanced to achieve less dropping rates. The study used the K-Means algorithm to group the learners according to reasons of success or failure while attending the MOOC. According to the 5M framework, five groups are identified: Management, Method, Material, Matter, and Measure. In general, some special patterns related to reasons for drop-out or successful completion were detected of the MOOCs. For example, the clusters "Management" and "Methodologies" were found highly related with learner dissatisfaction along with more attrition rate. Other clusters like "Materials" and "Machinery" were quite associated with high learner satisfaction along with course completion[11].

Another study in literature provides an overview of several data mining techniques used for predicting the performance of learners in MOOCs. Several machine learning algorithms and mining techniques for the prediction of MOOCs' students' performances were discussed, including XGBoost Algorithm, Graph Convolutional Networks (GCN), Logistic Regression, Random Forest, Support Vector Machine (SVM), as well as Deep Learning involving attributes such as Engagement Metrics, Demographics and Socio-Economic Features, Course Interactions, and Assessment Scores. This study puts a strong focus on the role of engagement measures and demographic factors in predicting success and emphasizes that such advanced machine learning techniques as Graph Convolutional Networks yield the best accuracy[12].

Wen Xiao et al. described the effective aspects that might influence online learning performance via the understanding of machine learning models. The authors discovered that among all datasets, online learning behaviour characteristics include the number of interactions in course materials, video viewing frequency, and discussion participation as the most influential regarding online learning success. Much lower impacts were found as compared to online learning behaviours in demographic and academic characteristics[13].

An article used machine learning algorithms on predicting academic performance and assessing the learning behaviours of a student in MOOCs. The researchers found that certain specific behaviours have a strong correlation to final examination scores. ‘Video Viewing Time’, ‘Number of chapters scores’, ‘Study discuss several’ and ‘Homework Scores’ were used in the study for their findings. For instance, a correlation of high degree existed between homework scores and final examination scores. Students who performed poorly in homework were likely to fail their finals[14].

A further contribution by Jingjing et al. introduces the concept of learning cycles and demonstrates its usage in order to analyse and explain the behaviours during MOOCs learning. From the results, the potential for desire and intensity toward predicting student performance is determined in MOOCs. The results of this work can be applied in order to support educators and platform developers on ways to enhance the interest of students and to diminish dropout rates in online classes[15].

MOOCs have, more than anything else, traditionally influenced education to a wide extent, especially on such accessible and flexible learning platforms around the globe. By democratizing millions of lessons free of charge from major institutions such as Coursera, edX, or Udacity [16], this has been consistent with broader idea of open education realizing the removal of barriers while improving learning inclusiveness.

Although MOOCs supplement traditional education with increasing student engagement and results, problems include a high rate of dropouts and an absence of personal contact. Students enjoy the flexibility to learn at their own speed and at their convenience but find it challenging to impose self-discipline in the completion of such courses. The lack of support mechanisms that a structured course presents may make it easier to disengage from, particularly for those students who could not sustain keeping up with the subject matter individually[17]​.

A blended learning paradigm that encompasses MOOCs within regular courses provides an answer because it integrates the strengths of both approaches. Studies conducted with students in Master's programs at Politehnica University of Timisoara have shown that such a model can achieve greater student satisfaction and improved educational results due to delivering the freedom of online learning while preserving structure and support[18].

With over thousands of students enrolling in universities daily, MOOCs have emerged as a transformative force in the Indian education system, enhancing accessibility and knowledge dissemination at a minimal cost. The three authors analyze India’s higher education challenges, such as faculty shortages and low per-student spending, and propose MOOCs as a viable solution. They explore major MOOC platforms, including edX, Coursera, and Udacity, along with India’s NPTEL initiative, to highlight their role in improving educational quality. Their research includes user perception analysis, revealing that most learners view MOOCs as a complement to traditional education, aiding in skill development and lifelong learning [19].

# Gap Analysis

MOOCs have been widely accepted for democratizing education, more research is required on how such courses affect student performance on different types of assessments. Our study looks to fill important lacunas within the existing literature on MOOCs by considering students' performance in both internal and external assessments within two different MOOCs

The majority of current MOOCs research focus on overall student success indicators or specific components, including participation rates, video engagement, and final exam results. For instance, Sahar Voghoei et al. and Tianping Deng et al. researched the relationships between online participation and general academic success, which emphasize the need for continued interaction.

These studies often fail to differentiate the forms of assessments that occur in the same MOOC. Internal assessments are often condensed into a single performance statistic or overshadowed by emphasis on final exams or other external evaluations.

The literature also handles student performance as a composite outcome without discussing the performance in internal assessments, which take place during the course and are integral to continuous learning, compares with performance in external assessments, which are often cumulative and take place at the end of the course.

Although the volume of studies on MOOCs is increasing, only a few studies explicitly compared the performance of students in internal versus external assessments based on the same MOOC. This distinction is important because internal assessments are designed to track and encourage continuous learning, whereas external exams often tests the final degree of understanding or recall of course material.

Our study will fill this gap by conducting a comprehensive investigation of how students perform in these two different types of exams. This will involve exploring the relationship between internal continuous formative assessments and external final summative exams performance. The results may indicate whether students who do well in continuing assessments are equally effective in final exams or if there is a discrepancy that calls for different instructional strategies or further help.

# Methodology

## Dataset Description

The dataset used to analyse student performance in technical and non-technical programs, as well as to perform clustering based on their marks in internal and proctored examinations to identify patterns and groups of students with similar academic performance, is made up of student performance data from a technical course called ‘Python for Data Science’ and a non-technical course named ‘Effective Writing’. Each record represents a student and consists of the different attributes resembling Unique id (A serial number assigned to each student), Student Name (The name of the student (anonymized in the dataset)), Program **(**The academic program the student is enrolled in (e.g., B.Sc. Biotechnology, B.Tech, M.Tech, etc)), Assignment Marks **(**The marks obtained by the student in internal assignments out of 25), Proctored Examination Marks (The marks obtained by the student in proctored examination out of 75), and Final Marks (The total marks obtained by the student, calculated as the sum of assignment marks and proctored examination marks, out of 100).

## Data Analysis

We had preprocessed data for the study to check the authenticity and reliability of data. We cleaned the data set from outliers that would skew the outcome.

After preprocessing, we ran K-Means clustering to analyze and compare academic performance of students between the technical and non-technical programs. K-Means clustering is an unsupervised form of machine learning, clustering data into distinct groups or clusters of similarity [9]. It iterates, assigns points to each closest centroid at a given point; then resets centroids to calculated places based upon points assigned. Thus, process iterates in rounds until cluster centroids become stable, implying also that resulting clusters well have defined meaning.

We divided learners into two categories: technical and non-technical. Technical programs characterize themselves by the emphasis placed upon the thorough understanding of certain specific technologies, programming languages, tools or techniques [20]. Non-technical programs, on the other hand, focus on an overall understanding of technology instead of a deep dive in the hands-on technical features. Instead, they might focus on management, strategy, communication, or even other areas of specialization [20]. We classified the students based on these differences that allow for a more subtle analysis of their academic performance in such diverse categories.

In this study, we used K-Means clustering to group the students according to their performance both on internal and proctored tests. The K-means clustering algorithm uses the cluster’s object mean value to generate clusters. The cluster number is required as a user parameter and is used in the arbitrary cluster center selection from the dataset in the standard K-means algorithm. The centroid vector C = { c1, c2, …, ck } can be easily recognized with a set of input data supplied to the K-means clustering algorithm, where K is the number of centroids defined by the user. The K-means clustering algorithm is appreciated due to its flexibility, efficiency, and ease of implementation. It is also among the top ten clustering algorithms in data mining. K-means clustering algorithm is widely accepted due to the simplicity and low computational complexity and serves in solving clustering problems in many domains. Several variants of K-means clustering have been developed to enhance its performance. In this study, we compared the clusters generated for technical students with that of non-technical students to visually understand trends and gain insights regarding how students from different types of academic backgrounds perform across different courses [21].

# Results

The two data sets, ‘Effective Writing’ and ‘Python for Data Science’ contained the result of internal and proctored test of students on which the study was conducted. The main objective of the study was to classify the programs of the students into "technical" and "non-technical" and to apply K-Means clustering to determine their performance patterns and compare the average scores between the two groups.

There were 4363 student records in the dataset for the course ‘Effective Writing’. Programs were categorized as technical and non-technical based on terms such as B.Tech, M.Tech, Computer, Engineering, and Information Technology. From this classification, it emerged that 656 students were in technical programs, and the rest 3707 were in non-technical programs.

After this categorization, the standardized internal and proctored scores were passed through K-Means clustering, and two different groupings emerged. The students were termed technically and non-technically classified based on their curricula and how their grades compared with internal and proctored grades.

Average internal and proctored scores were determined for each group. The average of technical students in internals was at 19.53; for non-technical it was at 18.82. Similarly, on the average proctored score for technical students stands at 50.22 while that of non-technical stands at 46.44. This research also points out to several students whose internal and proctored results exceeded category averages. Among the technical, 460 scored above than the average internal score but 326 scored above average proctored result. On the other hand, a much higher number of non-technical students appeared above average: 2685 in internals while 1993 in proctored.

As illustrated in figure 2, a scatter plot was created for demonstration of results showing clustering of the student based on internal and proctored scores. The plot is quite successful in establishing that technical students (blue) are easily differentiated from the non-technical ones (red) because each of the two groups occupies parts of scatter space. It shows clustering with a non-technical pupil dominating.

A diagram of red and blue dots

Description automatically generated

Figure 2 K-Means Clusters with Tech and Non-Tech Students for Effective Writing

The dataset for the course ‘Python for Data Science’ dataset contained 2321 student records. The programs were categorized as technical and non-technical based on terms such as B.Tech, M.Tech, Computer, Engineering, and Information Technology. Based on this categorization, 2214 students were enrolled in technical courses, and the remaining 107 students were on non-technical courses.

After stratification, K-Means clustering was done on standardized internal and proctored scores that resulted in two different clusters. The students were stratified into technical and non-technical students based on the curriculum taken, and the internal and proctored marks obtained were compared. Average internal and proctored scores for each category were calculated. For internal scores, technical students averaged at 21.9 while that of non-technical students averaged at 22.7.

Similarly, technical students averaged a proctored score of 43.09 while non-technical scored at 38.11. It is also found that there are students whose internal and proctored results exceed the averages of their respective categories. 1617 Technical students have above the average of internal score while 1080 students did in proctored mark. However, a considerably greater number of non-technical students performed better than the average where 83 were found above the average in internals and 54 in proctored.

As illustrated in figure 3, a scatter plot was generated to show the way students cluster on internal and proctored scores. This plot clearly demonstrated that there is a distinction between the Technical (blue) learners and the non-technical (red) learners; two groups occupy distinct parts of the scatter space. Such clustering shows technical learners as dominant.

A graph of blue and red dots

Description automatically generated

Figure 3 K-Means Clusters with Tech and Non-Tech Students for Python for Data Science

# Conclusion

The study was aimed at investigating the performance of students on internal and external assessments from two different technical and non-technical MOOCs. The continuous internal assessment and final proctored exam relationship carried new vital information through the use of K-Means clustering on the data of the student performance. The results showed that in technical subjects, like ‘Python for Data Science,’ a high correlation between internal and external assessment scores was mostly found. Thus, in proctored assessments, the highest measures of success for technical courses are indeed consistent performance and continuous participation during the course. In contrast, students in non-technical courses, like ‘Effective Writing,’ mostly have more variation in terms of performance. This variation emphasizes a need for varied teaching designs tailored to different learning approaches than present in non-technical classes.

The result emphasizes the importance of support networks and diversified teaching strategies in MOOCs in order to meet the diverse needs of students.

Although MOOCs have succeeded in being an innovative and democratizing way to enhance education, this research further opens the possibility that every subject or student population would not benefit equally in implementing a one-size-fits-all strategy. Teachers must instead consider creating more personalized paths of learning that help reach the unique needs and talents of students in both technical and non-technical subjects, maximizing learning outcomes. It is therefore crucial to keep on focusing on the continuous development of teaching strategies and assessment techniques as MOOCs continue to evolve and get more integrated into formal education systems. This research contributes to the ongoing debate by demonstrating how a better understanding of students' performance on various kinds of assessments may lead to better teaching and learning strategies.

Further research efforts would then be able to build upon these findings by exploring more auxiliary variables, such as profiles of students and engagement data, to further develop virtual learning methodology.

##### References

1. M. Khalil, J. Wong, B. De Koning, M. Ebner, and F. Paas, “Gamification in MOOCs: A review of the state of the art,” in IEEE Global Engineering Education Conference, EDUCON, IEEE Computer Society, May 2018, pp. 1629–1638. doi: 10.1109/EDUCON.2018.8363430.
2. Dan. Suthers, Proceedings of the Third International Conference on Learning Analytics and Knowledge. ACM, 2013.Jason Jerald. 2015. The VR Book: Human-Centered Design for Virtual Reality. Association for Computing Machinery and Morgan & Claypool.
3. K. Foon Hew, C. Qiao, and Y. Tang, “Understanding Student Engagement in Large-Scale Open Online Courses: A Machine Learning Facilitated Analysis of Student’s Reflections in 18 Highly Rated MOOCs,” 2018. [Online]. Available: <https://www.class-central.com/report/mooc-stats-2016/>
4. D. A. Joyner, “Meet Me in the Middle: Retention in a ‘MOOC-Based’ Degree Program,” in L@S 2022 - Proceedings of the 9th ACM Conference on Learning @ Scale, Association for Computing Machinery, Inc, Jun. 2022, pp. 82–92. doi: 10.1145/3491140.3528283.
5. I. Ahmad, S. Sharma, R. Singh, A. Gehlot, N. Priyadarshi, and B. Twala, “MOOC 5.0: A Roadmap to the Future of Learning,” Sep. 01, 2022, MDPI. doi: 10.3390/su141811199.
6. L. Yuan and S. Powell, “MOOCs and Open Education: Implications for Higher Education”, doi: 10.13140/2.1.5072.8320.
7. M. J. Page et al., “The PRISMA 2020 statement: An updated guideline for reporting systematic reviews,” Mar. 29, 2021, BMJ Publishing Group. doi: 10.1136/bmj.n71.
8. S. Voghoei, N. Hashemi Tonekaboni, D. Yazdansepas and H. R. Arabnia, "University Online Courses: Correlation between Students' Participation Rate and Academic Performance," 2019 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 2019, pp. 772-777, doi: 10.1109/CSCI49370.2019.00147.
9. T. Deng, L. Zhang, and X. Hei, “Analysis of Learning Behavior Based on MOOC Data,” in *TALE 2021 - IEEE International Conference on Engineering, Technology and Education, Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 94–98. doi: 10.1109/TALE52509.2021.9678539.
10. T. Phan, S. G. McNeil, and B. R. Robin, “Students’ patterns of engagement and course performance in a Massive Open Online Course,” *Comput Educ*, vol. 95, pp. 36–44, Apr. 2016, doi: 10.1016/j.compedu.2015.11.015.
11. S. Soukaina, S. El Miloud, and M. El Hassan Charaf, “MOOCs performance analysis based on quality and machine learning approaches,” in *2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science, ICECOCS 2020*, Institute of Electrical and Electronics Engineers Inc., Dec. 2020. doi: 10.1109/ICECOCS50124.2020.9314606.
12. N. Srivastava and J. Ahmad, “A Review on the Learner’s Performance Prediction Techniques in MOOC Courses through Data Mining,” in *5th IEEE International Conference on Advances in Science and Technology, ICAST 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 314–317. doi: 10.1109/ICAST55766.2022.10039570.
13. W. Xiao and J. Hu, “Analyzing Effective Factors of Online Learning Performance by Interpreting Machine Learning Models,” *IEEE Access*, vol. 11, pp. 132435–132447, 2023, doi: 10.1109/ACCESS.2023.3334915.
14. H. Guo, Y. Li, F. Liu, and W. Hu, “Machine-Learning based MOOC learning data analysis,” in Proceedings - 2021 7th IEEE International Conference on Big Data Security on Cloud, IEEE International Conference on High Performance and Smart Computing, and IEEE International Conference on Intelligent Data and Security, BigDataSecurity/HPSC/IDS 2021, Institute of Electrical and Electronics Engineers Inc., May 2021, pp. 63–68. doi: 10.1109/BigDataSecurityHPSCIDS52275.2021.00022.
15. J. He, C. Men, S. Fang, Z. Du, J. Liu, and M. Li, “Analysis of MOOC Learning Rhythms,” in Proceedings - 20th International Conference on High Performance Computing and Communications, 16th International Conference on Smart City and 4th International Conference on Data Science and Systems, HPCC/SmartCity/DSS 2018, Institute of Electrical and Electronics Engineers Inc., Jan. 2019, pp. 1555–1562. doi: 10.1109/HPCC/SmartCity/DSS.2018.00255.
16. R. Al-Shabandar, A. J. Hussain, P. Liatsis, and R. Keight, “Analyzing Learners Behavior in MOOCs: An Examination of Performance and Motivation Using a Data-Driven Approach,” *IEEE Access*, vol. 6, pp. 73669–73685, 2018, doi: 10.1109/ACCESS.2018.2876755.
17. T.-M. Rhyne, H. Qu, and Q. Chen, “Visualization Viewpoints Visual Analytics for MOOC Data,” 2015.
18. S. Parvathavarthini, S. Sharvanthika K, M. Jagadeesh, and B. Kishore, “Analysis of Student Performance in E-learning Environment using Crow search based Fuzzy clustering,” in *Proceedings - 2nd International Conference on Smart Electronics and Communication, ICOSEC 2021*, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 1784–1787. doi: 10.1109/ICOSEC51865.2021.9591920.
19. P. K. Singh, I. Gandhi and P. Nand, "MOOCs: The paradigm-shift in Indian education," 2014 IEEE International Conference on MOOC, Innovation and Technology in Education (MITE), Patiala, India, 2014, pp. 317-320, doi: 10.1109/MITE.2014.7020295.
20. M. Wang, “Hybrid data clustering algorithm and interactive experience in E-learning electronic course simulation of legal education,” *Entertain Comput*, vol. 52, Jan. 2025, doi: 10.1016/j.entcom.2024.100760.
21. Abiodun M. Ikotun, Absalom E. Ezugwu, Laith Abualigah, Belal Abuhaija, Jia Heming, “K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data”, Information Sciences, Volume 622, 2023, Pages 178-210, ISSN 0020-0255, https://doi.org/10.1016/j.ins.2022.11.139.