# A Systematic Review of Predictive Factors for Learner Attrition in Online Learning: Insights for Machine Learning Models

Stanley Munga Ngigi
School of Pure and Applied Sciences
Kirinyaga University, Kutus, Kenya
*Email: sngigi [AT] kyu.ac.ke*

Dr. James Mwikya
School of Pure and Applied Sciences
Kirinyaga University, Kutus, Kenya
*Email jmwikya [AT] kyu.ac.ke*

Dr. Victor Mageto
School of Pure and Applied Sciences
Kirinyaga University, Kutus, Kenya
*Email: vmageto [AT] kyu.ac.ke*

*Abstract*----- **Over the past ten years, online education has expanded rapidly due to its accessibility, scalability, and flexibility. Despite its potential, high attrition rates in online education threaten both student progress and the legitimacy of the institution. A comprehensive analysis of empirical research on the factors influencing learner attrition in online learning settings is presented in this study. To identify the individual, course-level, institutional, and technical causes of attrition, it incorporates and categories the body of existing work. The results point to the complex aetiology of attrition and identify important domains for focused intervention and predictive modelling.**

***Key words: Learner Attrition, online learning, dropout, e-learning retention***

## I. INTRODUCTION

In the current digital era, online education has emerged as a crucial means of instruction, offering students worldwide opportunities that are both flexible and expandable (Hanna, 2019). Particularly during the COVID-19 epidemic years, the pace of digital education demonstrated both its transformative potential and its dangers (UNESCO, 2021). The effectiveness and validity of online education systems are severely hampered by high learner attrition rates, despite the fact that digital learning has increased. According to research, online learning dropout rates are 10–20% greater than in-person learning (Hachey et al., 2023), endangering both the viability of the institution and the success of its students.

A distinct set of technological, socioeconomic, and infrastructure limitations have influenced the uptake of online education in sub-Saharan Africa, and Kenya in particular (Internet World Stats, 2021). Even though Kenya's government has prioritized digital transformation in education through initiatives like Vision 2030, issues including poor internet connectivity, limited device access, and low levels of digital literacy have continued to contribute to attrition. Only 50% of Kenyan university students regularly had access to reliable internet during the COVID-19 lockdown, according to a poll conducted by the Commission for University Education (CUE). This led to a significant rise in online course dropout rates.

Predicting learner attrition accurately is essential for timely intervention. Early identification of high-risk students can improve academic performance, lower dropout rates, and protect revenue sources for the school (Adnan et al., 2021). By offering individualized learning routes and focused coaching, predictive models can optimize support and engagement from the standpoint of the student. (Kok et al., 2024) In light of this, machine learning has become a leading contender for attrition prediction by analyzing large and intricate data sets, such as academic performance, learner demographics, and interaction patterns.

However, it is crucial to understand the reasons behind learner attrition based on a comprehensive analysis of research findings before developing predictive models that work. A growing corpus of research examines a wide range of academic, psychological, institutional, and technological aspects that influence students' decision to discontinue online courses (Raman et al., 2021) . However, findings are often dispersed, contextually specific, and not framed by a shared paradigm.

This highlights the necessity of doing a systematic review that integrates and synthesises these findings to get a more comprehensive understanding of the factors that influence attrition.

This review's objective is to critically evaluate peer-reviewed research in order to determine the most often mentioned causes of learner attrition in online learning environments. This study offers fundamental information to guide the development of data-based prediction models and intervention strategies by classifying these into thematic groups. Additionally, it provides evidence-based suggestions on learner retention that impact policy and practice, especially in low-resource contexts like Kenya.

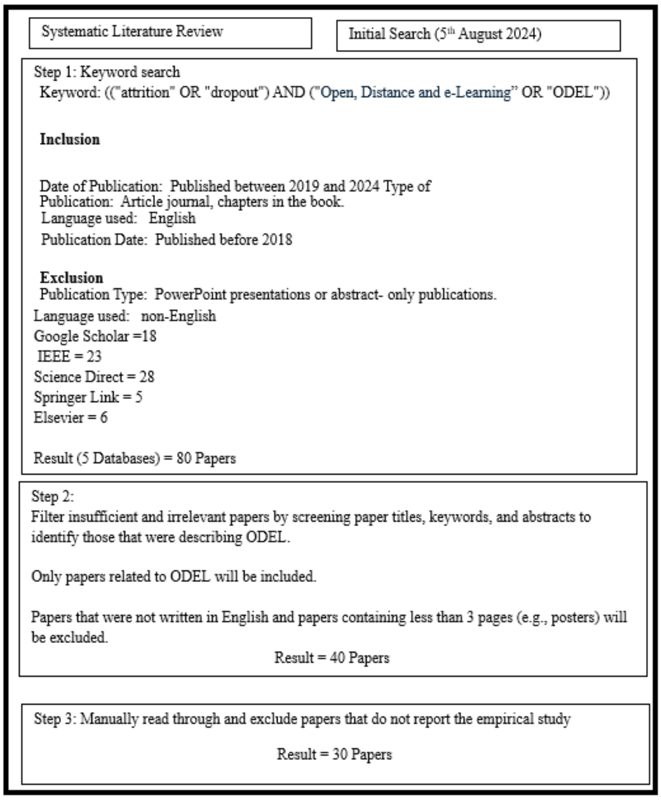
## II. LITERATURE REVIEW

A number of peer-reviewed studies on learner attrition in e-learning that were published between 2019 and 2024 were retrieved from scholarly databases. According to existing research, the sudden switch to online learning, particularly in the wake of the COVID-19 epidemic, revealed flaws in learner engagement, preparedness, and perseverance.

Even while online platforms allowed for learning continuity, the majority of students experienced problems like digital exhaustion, poor connectivity, and a lack of emotional support, which raised attrition rates (Rahmani et al., 2024). Interestingly, little research has been done on using predictive models that include institutional, behavioral, and demographic data to identify students who are at risk (Hung et al., 2019). This indicates a lack of progress in creating suitable machine learning models, including ensemble models, that can aid in attrition prevention and early intervention in online learning environments.

## III. METHODOLOGY

In order to determine the main factors influencing learner attrition in online learning settings, this study used a methodical literature review technique. Google Scholar, IEEE Xplore, Science Direct, Springer Link, and Elsevier were the five academic databases used for the review, which focused on peer-reviewed literature published between 2019 and 2024.The search was guided by the keywords: (\"attrition\" OR \"dropout\") AND (\"online learning\" OR \"e-learning\" OR \"distance education\").




Learner attrition is greatly influenced by demographic factors in online university programs, including age, gender, geography, socioeconomic status, and employment status (Packham et al., 2004). Older students tend to have more responsibilities, including those pertaining to their workplaces or families, which may constrain them from continuing in online courses (Xavier & Meneses, 2022). Gender differences also come into play. Research has shown that female students can have certain issues with time management and juggling multiple tasks, which can impact their engagement and (Farrell & Brunton, 2020). Higher dropout rates can also be triggered by students from different geographical areas having different levels of access to required resources (Bawa, 2016). Another important consideration is socioeconomic status; students from lower-income families frequently experience financial hardship or do not have the technology to continue, resulting in increased attrition (Petro et al., 2020). Part-time or full-time student status affects time management and commitment to program in students, which can result in increased dropout (Levy, 2021).

### *4.2 **Academic Performance***

Academic performance is one good indicator of attrition in online learning environments. A student's readiness for and potential success with online learning are at times determined by their prior academic achievement, i.e., their GPA or grades in the relevant subject matter areas (Kim et al., 2020). Current academic achievement, e.g., grades and course progress, also provide useful data with respect to a student's motivation and threat of attrition (Caruth, 2018). Students' engagement and commitment are demonstrated by their participation in assessment like quizzes, tests, and assignments, which has a direct impact on their likelihood to continue with the material (Chapman & Andrade, 2024). Further, achievement and persistence in distance learning depend heavily on students' capacity to cope with independent and asynchronous learning habits (Errabo et al., 2024).

### *4.3 **Engagement Metrics***

Student engagement indicators, including frequency of course access, interaction with course content, discussion forum participation, communication with instructors and peers, and use of support services, are essential indicators of learner retention in online learning. High LMS login frequency is associated with high engagement and low attrition rates (Talebi et al., 2024). Similarly, active participation in course materials such as videos, readings, and assignments is a measure of student engagement and can also be employed in predicting attrition risk (Villegas-Ch et al., 2024). Participating in online discussion forums facilitates a sense of belonging and community, which is critical in the prevention of dropout rates (Conceição & Biniecki, 2024). Active engagement with instructors and peers, including timely response and participation in group assignments, enhances a student's belongingness and resilience (Ravishankar et al., 2024). Lastly, the utilization of support services, e.g., tutoring or academic advising, is also positively related to student persistence and success in online courses (Dunlap, 2024).

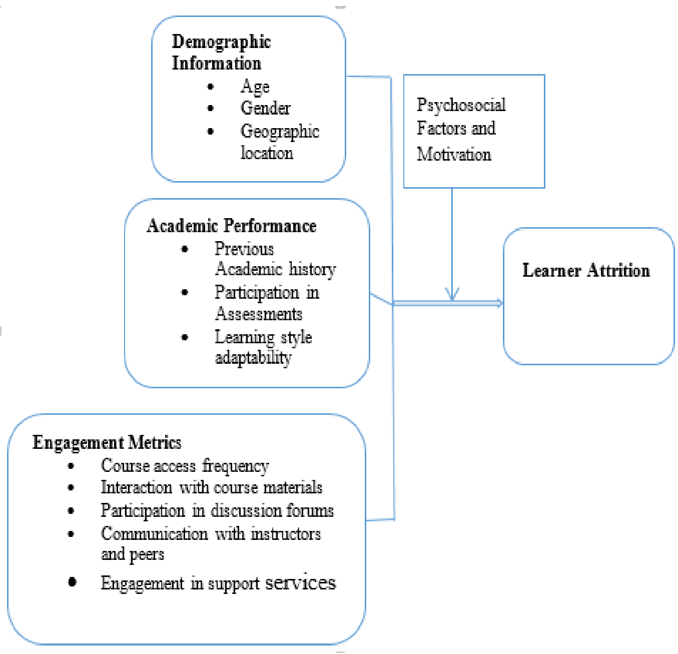



**Fig 2: Conceptual Framework of Factors Influencing learner Attrition in Online Learning**

The conceptual framework was created to highlight the most important elements influencing student attrition in online learning based on a systematic study of the literature. Students' persistence in online programs is first and foremost influenced by demographic parameters such as age, gender, socioeconomic position, occupation, and geographic location. Time and technology availability issues are common for older students, students from underprivileged neighborhoods, and students from low-income households.

Second, learning performance is an excellent indicator of attrition; a higher likelihood of leaving is predicted by bad grades, little advancement, or inadequate preparation. Last but not least, engagement metrics including frequency of logins, engagement with the course material, discussion participation, communication with instructors, and assistance utilisation are important markers of student commitment. Attrition is often preceded by low levels of participation.

These combined factors serve as the basis for the prediction model used in this study, which identifies susceptible learners and suggests early treatments using ensemble models, particularly gradient boosting and neural networks.

| **Title of Paper** | **Theme of Research** | **Year of Study** | **Country of origin** |
|---|---|---|---|
| Factors that Contribute to Students' Attrition in Open and Distance Learning (ODL) Environment | Challenges of OLD learning environments. | 2022 | Malaysia |
| Factors affecting student dropout in MOOCs | Understanding causes of dropout in MOOCs | 2019 | South Africa |
| Prediction of students' early dropout based on their interaction logs in online learning environment | Using interaction logs to predict early dropouts in online education | 2019 | China |
| Factors Contributing to Student Retention in Online Learning and Recommended Strategies for Improvement: A Systematic Literature Review | factors that affect online learning completion and retention. | 2019 | USA, Virginia |
| Factors Affecting Students' Preferences for Online and Blended Learning: Motivational vs. Cognitive | Learner preferences impacted by cognitive and motivational elements | 2022 | Europe |

| | | | |
|---|---|---|---|
| Persistence and Dropout in Higher Online Education | Dynamics of persistence and dropout in higher education | 2022 | Pakistan |
| Factors affecting students' intentions to undertake online learning: an empirical study in Vietnam | In online learning, learner motivation and intention | 2021 | Vietnam |
| Predicting Dropout in Online Learning Environments | Predictive modeling of learner dropout | 2023 | UAE |
| Predicting student dropout in subscription-based online learning environments: The beneficial impact of the logit leaf model | Predicting dropout in subscription-based learning using AI models | 2021 | France |
| Dropout management in online learning systems | Techniques for reducing dropouts in online education | 2021 | India |
| Psychosocial Implications, Students Integration/Attrition, and Online Teaching and Learning in South Africa's Higher Education Institutions in the Context of COVID-19 | Structure and psychological issues in online learning during COVID-19 | 2023 | South Africa |
| Dropout in online higher education | Analysis of online higher education dropout rates | 2024 | Iran |
| The effectiveness of e-tutoring in an open and distance e-learning environment: evidence from the university of south Africa | E-tutoring's effectiveness in ODL settings | 2021 | South Africa |
| Investigating the E-Learning Challenges Faced by Students during COVID-19 in Namibia | Online learning obstacles during COVID-19 | 2021 | Namibia |
| The MOOC dropout phenomenon and retention strategies | MOOC dropout trends and retention tactics | 2021 | China |
| Rurality and Dropout in Virtual Higher Education Programmes in Colombia | Rural context's effect on dropout rates from virtual schooling | 2022 | Colombia |
| Charting the Course of School Dropout Research: A Bibliometric Exploration | Bibliometric examination of dropout research trends and gaps | 2024 | Romania |
| The effectiveness and efficiency of student support services in open distance learning institutions in Africa | Assessment of student assistance programs in ODL | 2021 | South Africa |

| | | | |
|---|---|---|---|
| | institutions in Africa | | |
| Online Learning in Higher Education during COVID-19 Pandemic: A case of Ghana | Institutional reaction to COVID-19's online education | 2021 | Ghana |
| Learner Dropout in South African Schools | Examining the reasons behind school dropouts | 2021 | South Africa |
| Risk factors associated with first-year students' intention to drop out from a university in South Africa | Risk factors for first-year university students' intention to drop out | 2020 | South Africa |
| Persistence in a Game-Based Learning Environment: The Case of Elementary School Students Learning Computational Thinking | Persistence in computational learning environments based on games | 2020 | Europe |
| School–community interventions to curb learner dropout: The perceptions of key education stakeholders in a rural South African school neighborhood | Community-school projects' contribution to lower learner dropout rates | 2021 | South Africa |
| Student Support as a Panacea for Enhancing Student Success in an Open | Support from students as a tactic to | 2023 | United States |

| | | | |
|---|---|---|---|
| Distance Learning Environment | increase ODL success | | |
| Course satisfaction and perceived learning among distance learners in Malaysian Research Universities | Learning outcomes and course satisfaction in Malaysian ODL | 2024 | Malaysia |
| Exploring Factors, and Indicators for Measuring Students' Sustainable Engagement in e-Learning | Metrics and indicators of long-term e-learning student involvement | 2019 | Korea |
| Management of School Environmental Factors on Dropout Rates on Public Primary Schools in Kuresoi South Sub County, Kenya | Environmental factors' impact on elementary school dropout rates | 2021 | Kenya |
| Social Media as a Determinant of Students' Dropout Rates in Secondary Schools in Kenya | Social media's effect on dropout rates in secondary schools | 2023 | Kenya |
| Obstacles to successful uptake of open distance and e-learning (odel) programmes: a case of kenyatta university, Kenya | Obstacles to Kenyan ODeL program enrolment and completion | 2022 | Kenya |
| | | | |

## V. CONCLUSION AND RECOMMENDATIONS

In the current day, online learning has become a crucial part of education and will only grow in various learning settings. It is imperative to address the ongoing issue of learner attrition as more and more schools embrace online modalities. In addition to impeding student achievement, high turnover rates also undermine the sustainability and legitimacy of online learning platforms. Therefore, in order to ensure high-quality and equitable learning outcomes, institutions have a positive obligation to create ways to anticipate and avoid attrition.

Learner engagement, instructional design, technological access, institution support, and individual learner characteristics are some of the determinant factors for learner attrition in online learning environments that were developed by this systematic review. These factors attest to the intricacy of attrition and the requirement for dimensional solutions.

It is recommended that a case study of contemporary online students be used to apply and test the conceptual framework that was developed from this review. The framework will be improved by empirical validation, which will also enable the creation of prediction models that higher education institutions can use to spot potentially vulnerable students early on and implement appropriate interventions.

## REFERENCES

[1] Adnan, M., Habib, A., Ashraf, J., Mussadiq, S., Raza, A. A., Abid, M., Bashir, M., & Khan, S. U. (2021). Predicting at-Risk Students at Different Percentages of Course Length for Early Intervention Using Machine Learning Models. *IEEE Access*, *9*, 7519–7539. https://doi.org/10.1109/ACCESS.2021.3049446

[2] Bawa, P. (2016). Retention in Online Courses: Exploring Issues and Solutions—A Literature Review. *Sage Open*, *6*(1), 2158244015621777. https://doi.org/10.1177/2158244015621777

[3] Caruth, G. D. (2018). Student Engagement, Retention, and Motivation: Assessing Academic Success in Today's College Students. *Participatory Educational Research*, *5*(1), Article 1. https://doi.org/10.17275/per.18.4.5.1

[4] Chapman, J. R., & Andrade, M. (2024). Improving part-time instructors' student failure rate with an educational engagement information system. *Educational Technology Research and Development*, *72*(3), 1465–1482. https://doi.org/10.1007/s11423-024-10352-2

[5] Conceição, S. C. O., & Biniecki, S. M. Y. (2024). Closed and Open Online Discussion Forums. In *Methods for Facilitating Adult Learning*. Routledge.

[6] Dunlap, K. (2024). The impact of academic advising on persistence in nontraditional students completing a baccalaureate degree online. *Theses and Dissertations*. https://scholarsjunction.msstate.edu/td/6101

[7] Errabo, D. D., Dela Rosa, A., & Gonzales, L. J. M. (2024). Optimizing differentiated podcasts to promote students' self-regulation and engagement, self-efficacy and performance in asynchronous learning. *Journal of Research in Innovative Teaching & Learning*, *17*(2), 368–390. https://doi.org/10.1108/JRIT-02-2024-0039

[8] Farrell, O., & Brunton, J. (2020). A balancing act: A window into online student engagement experiences. *International Journal of Educational Technology in Higher Education*, *17*(1), 25. https://doi.org/10.1186/s41239-020-00199-x

[9] Hanna, D. E. (2019). Higher Education in an Era of Digital Competition: Emerging Organizational Models. *Online Learning*, *2*(1). https://doi.org/10.24059/olj.v2i1.1930

[10] Hung, J.-L., Shelton, B. E., Yang, J., & Du, X. (2019). Improving Predictive Modeling for At-Risk Student Identification: A Multistage Approach. *IEEE Transactions on Learning Technologies*, *12*(2), 148–157. https://doi.org/10.1109/TLT.2019.2911072

[11] Kim, D., Lee, Y., Leite, W. L., & Huggins-Manley, A. C. (2020). Exploring student and teacher usage patterns associated with student attrition in an open educational resource-supported online learning platform. *Computers & Education*, *156*, 103961. https://doi.org/10.1016/j.compedu.2020.103961

[12] Kok, C. L., Ho, C. K., Chen, L., Koh, Y. Y., & Tian, B. (2024). A Novel Predictive Modeling for Student Attrition Utilizing Machine Learning and Sustainable Big Data Analytics. *Applied Sciences*, *14*(21), Article 21. https://doi.org/10.3390/app14219633

[13] Packham, G., Jones, P., Miller, C., & Thomas, B. (2004). E-learning and retention: Key factors influencing student withdrawal. *Education + Training*, *46*(6/7), 335–342. https://doi.org/10.1108/00400910410555240

[14] Petro, G., Gonzales, A., & Calarco, J. (2020). "Out of Luck": Socio-Economic Differences in Student Coping Responses to Technology Problems. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–10. https://doi.org/10.1145/3313831.3376156

[15] Rahmani, A. M., Groot, W., & Rahmani, H. (2024). Dropout in online higher education: A systematic literature review. *International Journal of Educational Technology in Higher Education*, *21*(1), 19. https://doi.org/10.1186/s41239-024-00450-9

[16] Raman, R., B, S., G, V., Vachharajani, H., & Nedungadi, P. (2021). Adoption of online proctored examinations by university students during COVID-19: Innovation diffusion study. *Education and Information Technologies*, *26*(6), 7339–7358. https://doi.org/10.1007/s10639-021-10581-5

[17] Ravishankar, S., Spencer-Drakes, T. C. J., Fernandes, I. H., Hayes, M. I., Coopwood, S., Spencer, I., & Neal, S. E. (2024). Empowering STEM students: A university-wide mentorship program fostering retention and belonging. *Journal of Cellular Physiology*, *239*(7), e31348. https://doi.org/10.1002/jcp.31348

[18] Talebi, K., Torabi, Z., & Daneshpour, N. (2024). Ensemble models based on CNN and LSTM for dropout prediction in MOOC. *Expert Systems with Applications*, *235*, 121187. https://doi.org/10.1016/j.eswa.2023.121187

[19] Villegas-Ch, W., García-Ortiz, J., & Sánchez-Viteri, S. (2024). Application of Artificial Intelligence in Online Education: Influence of Student Participation on Academic Retention in Virtual Courses. *IEEE Access*, *12*, 73045–73065. IEEE Access. https://doi.org/10.1109/ACCESS.2024.3403758

[20] Xavier, M., & Meneses, J. (2022). Persistence and time challenges in an open online university: A case study of the experiences of first-year learners. *International Journal of Educational Technology in Higher Education*, *19*(1), 31. https://doi.org/10.1186/s41239-022-00338-6