African Journal of Education Science and Technology (AJEST) Vol. 8 No. 4 (May, 2026)
University of Eldoret, Kenya, Mount Kenya University, Kenya, Chukwuemeka Odumegwu Ojukwu University, Nigeria, Kyambogo University, Uganda and University of Makeni, Sierra Leone

Explainable Deep Learning for Predicting Student Dropout in Kenyan Universities: A Systematic Review of Models, Risk Factors, and Interpretability

***Bonfas Ogolla [1], Argan Wekesa[2], and Casper Shikali[3]**

[1]*Department of Computer Science and Information Technology, Co-operative University of Kenya,*
[2] *Department of Mathematical Sciences, Co-operative of Kenya,*
[3]*Department of Computer Science, South Eastern University Of Kenya.*

***Corresponding Email:** ogolla25.bonfas@student.cuk.ac.ke

Abstract

Student dropout continues to hinder higher education in Kenya, and one of the main defenses against it is identifying at-risk students early. Deep learning outperforms conventional machine learning in predictive power, yet universities remain cautious about adopting it because its decisions are hard to explain. Explainable artificial intelligence tools such as SHAP and LIME are meant to open up these black-box predictions for university officials, but whether they have actually been applied to dropout prediction in Kenya or elsewhere in Africa is still unclear. This review gathered and assessed the available evidence on explainable deep learning for student dropout prediction, with particular attention to what it means for Kenyan universities. Eleven keywords guided the search, returning 1,628 candidate records. After removing duplicates, screening titles and abstracts, and checking eligibility, 26 papers were retained for evaluation. Risk of bias was assessed with PROBAST, and results were synthesized narratively around four research questions covering predictor factors, deep learning architectures, explainable AI application, and relevance to Kenya and Africa. Academic and behavioral variables emerged as the most consistently reported predictors of dropout. Hybrid architectures that pair a deep feature extractor with a second-stage classifier delivered the strongest predictive performance whenever they were compared directly against single-architecture models. Only five of the 26 studies applied SHAP or LIME, and only two studies were conducted in Africa, one of them in Kenya. Deep learning has built a fairly solid evidence base for dropout prediction, but interpretability has not kept up with how complex these models have become, and evidence from Kenya remains scarce. This gap points to both the research that is still needed and the design choices that will shape a planned study built around an explainable deep neural network.

Keywords: *Deep learning, Explainable artificial intelligence, SHAP, Student dropout, Kenya*

Introduction

Universities across the globe spend significant amounts of money to attract students, develop academic programs and provide support services, yet many students fail to finish their programs of study. Students who quit school before completing their education are typically at a disadvantage position in terms of future earnings and job prospects compared with graduates (Oreopoulos & Petronijevic, 2013) and some of these students incur education-related debts without receiving testimonial. Colleges lose tuition revenue as well as faculty time spent on students who do not graduate (Loritee et al., 2026), which contributes to the overall slowdown of skilled labor supply that economies need to grow (Duro et al., 2026).

Many universities address dropout issues only after a student has already dropped out, with advisers and support personnel being called in once trends in poor academic performance, low attendance, unpaid tuition and the like have been observed (Tinto, 2006). The fields of Educational Data Mining and Learning Analytics have paved the way for early identification instead (Baker & Yacef, 2009), since Student Information Systems, Learning Management Systems, and financial and academic records produce considerable amounts of data on students' actions, results and engagement (Romero & Ventura, 2020).

Machine learning algorithms have achieved moderate to strong prediction performance in dropout studies (Chung & Lee, 2019; Sarker, 2021; Marcolino et al., 2025). However, these older methods still require manual feature engineering, which is labor-intensive and can overlook patterns an automated method would find (LeCun et al., 2015) and they inadequately capture the nonlinear ways in which student characteristics combine to produce dropout.

Deep learning builds its own layered representation of information instead of relying on manually selected features (LeCun et al., 2015) and it has produced strong results in educational prediction (Altabrawee et al., 2019; Sarker, 2021). However, the flexibility of deep learning methods implies some drawbacks; in particular, it should be taken into account that these techniques are frequently referred to as black boxes since the model can predict the chances of a student to drop out without explaining why a particular student has been flagged in a certain way (Adadi & Berrada, 2018). In this context, Explainable Artificial Intelligence (XAI) introduces a theory for making obscure outputs understandable for their users (Gunning, 2017). The SHAP was created by Lundberg and Lee (2017). LIME is offered by Ribeiro et al. (2016). Only a few recent publications refer to XAI together with deep learning in the sphere of predicting dropout and their evidence is still scattered (Bettahi et al., 2025).

Over the last twenty years, Kenyan institutions of higher learning have expanded greatly (Commission for University Education, 2024); however, graduation rates haven't kept pace with enrollment figures. Socioeconomic pressures affect some students more than others and new students encounter difficulties in their first year due to inadequate guidance and support (Ogawa, 2024; Stofile & Enwereji, 2025; Loritee et al., 2026). Kenyan universities have access to such data (Nyawira et al., 2025), but they rarely make use of this information for predictive purposes. A small number of Kenyan studies have applied machine learning to dropout (Mduma & Machuve, 2021; Nyawira et al., 2025), but none so far has examined deep learning or explainable AI for this purpose in Kenyan setting.

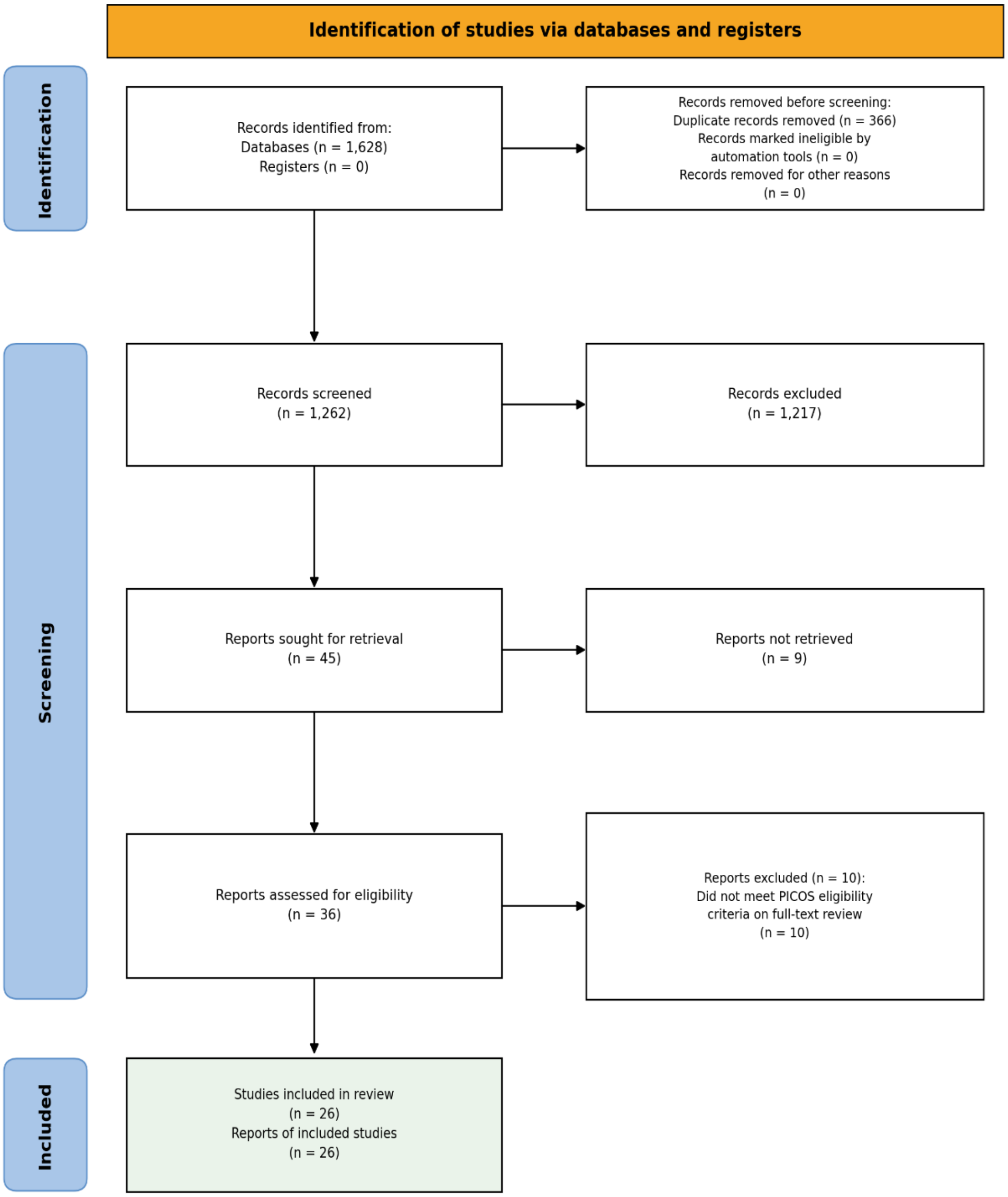
Two existing reviews come closest to this territory. Lopez-Muñoz et al. (2025) synthesized 75 machine learning studies on undergraduate dropout prediction, and Alharbi et al. (2026) reviewed SHAP-based interpretability in educational machine learning. Neither addresses deep learning and explainability together, and neither gives sustained attention to Kenya or wider Africa. Duro et al. (2026) cover both machine learning and deep learning approaches to dropout with attention to interpretability, deployment and ethics, but do not treat SHAP and LIME as a unit of analysis and say nothing about Kenya. Albugami et al. (2024) discuss explainability across a small set of studies without running a documented search and screening process.

This review addresses three gaps: focus on accuracy over interpretability, concentration of evidence outside Africa, and the absence of any review treating deep learning and XAI together for dropout prediction. This review synthesizes the evidence on explainable deep learning for student dropout prediction, with four research questions covering: (RQ1) predictor factors, (RQ2) deep learning architectures and performance, (RQ3) application of SHAP and LIME, and (RQ4) methodological quality and applicability to Kenya and Africa.

Materials and Methods

This review follows Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement (Page et al., 2021). Eligibility was assessed using PICOS criteria, adapted for a review of modeling techniques rather than clinical interventions. Thus, eligible studies must deal with some postsecondary students, implement some deep learning algorithm (one or more of following techniques: DNN, CNN, RNN, LSTM, GRU, transformer, or autoencoder) in predicting dropout, attrition, withdrawal, or any other type of non-completion outcome that would allow for comparing outcomes of different studies and provide at least one of the following results: accuracy, precision, recall, F1 score, or AUC-ROC. Exclusions are made for studies using only non-deep machine learning techniques without any deep learning algorithm implemented, and studies that predict other outcomes not related to dropout, such as scores or involvement of students. Also excluded are qualitative studies, publications in languages other than English without reliable translation, as well as thesis and non-peer-reviewed literature. Eligible publications span 1 January 2015 to the search date of August 2026, since 2015 marks the point at which deep learning began to see routine application in this field (LeCun et al., 2015).

Seven databases were searched between 1 and 3 August 2026: IEEE Xplore, ScienceDirect, ERIC, SpringerLink, ACM Digital Library, Google Scholar, and African Journals Online, the last included specifically to counter the bias toward North American and European publishers noted in the literature (Nyawira et al., 2025). A three-part search structure, a Core Search combining dropout and deep learning terms, an XAI Focus search adding explainability terms, and an African Context search adding regional terms, was run on every database using Boolean operators, phrase searching, and truncation adapted to each platform's syntax, producing twenty-one searches in total.



Screening proceeded in two stages, title and abstract followed by full text, with the author screening every record and the supervising academic acting as a second, independent reviewer, concentrated on uncertain or borderline records at the first stage and applied to all records at full-text stage. Disagreements were resolved by discussion against eligibility criteria, with referral to a third reviewer available though not required. Screening was managed in Rayyan (Ouzzani et al., 2016). Of 1,628 records identified across all databases, 366 duplicates were removed, 1,217 of the remaining 1,262 records were excluded at title and abstract screening, and 45 reports advanced to full-text retrieval. Nine could not be retrieved, 10 of the remaining 36 were excluded once the complete text could be checked against the eligibility criteria, and 26 studies were confirmed for inclusion.

Data were extracted using a CHARMS-based form (Moons et al., 2014) and piloted on three studies. The author extracted data from all 26 studies and the supervisor independently re-extracted a randomly selected third of them as a consistency check, with any discrepancy resolved by returning to the source study. Risk of bias was assessed independently by both reviewers using PROBAST (Wolff et al., 2019), which rates four domains: analysis, outcome, participants and predictors as low, unclear, or high risk.

The synthesis was based on a narrative and thematic approach due to the differences between the 26 studies, and it was created using a four-factor model (academic, demographic, socioeconomic, and institutional/behavioral) category according to SWiM guidelines. A quantitative meta-analysis of AUC-ROC using the valmeta function in the metamisc R package (Debray et al., 2017) was planned conditional on at least five studies reporting AUC-ROC with a usable measure of precision on a comparable outcome, and as a sensitivity check the narrative synthesis was repeated excluding studies rated high risk of bias on the PROBAST analysis domain. Certainty in the evidence for each research question was rated qualitatively as high, medium, low, or very low, adapting the logic of the GRADE approach (Guyatt et al., 2011) to a body of evidence with no intervention or comparison group to grade.

The studies that fall under the same architectural families were organized into four categories. Out of the 26 studies, nine employed hybrid architectures that make use of deep feature extractors and a second classifier, a nearly equal number of studies employed recurrent architectures including LSTM and GRU types, less numerous were those studying the convolutional architectures while the smallest number of studies utilized transformers, TabNet or auto-encoders as the main components.

Results and Discussion

Seven databases yielded 1,628 records, of which 26 studies met eligibility criteria after deduplication and two-stage screening. The included studies were published between 2017 and 2026, with 18 of the 26 (nearly 70%) appearing in 2025 or 2026, which suggests that explainable deep learning for dropout prediction is a young area of study rather than an established one. Nineteen of the 26 were conference papers, while seven were journal articles. As far as their location was concerned, evidence was mainly from Asia, since 10 studies (38%) were carried out in India, and the rest come from Bangladesh, Taiwan, Saudi Arabia, Jordan, UAE, Sri Lanka, and Vietnam. In Africa, there were only two studies, namely Kendagor et al. (2023), which was carried out in Kenya and the study by Hamal et al. (2025), which used two databases in Morocco. Also, many of the studies used same public benchmark datasets, instead of using new data collection, since the dataset "Predict Students' Dropout and Academic Success" from Portugal and the dataset OULAD from the UK were utilized in at least eight studies out of the 26.

The risk of bias evaluated using the four domains of PROBAST, comprising analyses, outcomes, participants and predictors and concluded that the weakest of these domains is the outcomes domain. Out of the studies reviewed, 20 out of 26 received a rating of unclear in the outcomes domain and only one was deemed to be of a high risk level. Most of the studies defined dropout simply based on one single dataset label and failed to mention what non-return interval was applied or how genuine withdrawal was differentiated from either a deferral or a transfer. The analysis domain, which checks for an independent test set, appropriate handling of class imbalance, and suitable performance metrics, showed 10 studies at low risk, 13 unclear, and 3 high risk. Three studies: Baranyi et al. (2020), Coppo et al. (2022) and Malkanthi et al. (2025), were rated low risk of bias across every domain, while four studies carried a high-risk rating on at least one domain.

Predictor Factors Associated with Dropout (RQ1)

Fourteen of the 26 studies reported predictor-level detail precise enough to classify into academic, demographic, socioeconomic, and institutional or behavioral categories. Academic predictors, prior grades, credit accumulation, assessment scores and attendance, appear most consistently, showing up in twelve of the fourteen studies with predictor-level reporting. Baranyi et al. (2020) SHAP-based ranking places program-specific academic performance among the twelve most influential features, while Hamal et al. (2025) found GPA and attendance rate among the strongest drivers of DeepDropNet's predictions using SHAP and Grad-CAM together. Institutional and behavioral predictors, platform engagement, submission timeliness and login frequency, form the second most consistent category, particularly in studies built on OULAD-style interaction logs, where Sivasankar et al. (2026) build an entire architecture, TERGATE-Warn, around treating weekly engagement as a sequence rather than a static count. It seems that there is an inconsistent presence of demographic and socioeconomic predictors. In the research conducted by Rajput et al. (2025), it is mentioned that predictors contributing to their model include age, previous education and financial status; in addition to this, the future solution was influenced by SHAP results mentioning the combination of debt and parental education as one of the biggest single factors identified during the research. Out of 26 studies analyzed, resource constraints were identified as predictors in only two studies, conducted in Kenya and Morocco, which makes it impossible to confirm that the findings can be applied to Kenya context. Nagy and Molontay (2024) note intervention-focused models favor academic and behavioral signals since these can be tracked repeatedly unlike fixed demographic traits. The near absence of resource-constraint predictors here may reflect that same bias and not just missing data. The pattern can be connected with Tinto's (2017) theory of student integration where the academic integration (the accomplishment of good grades, accumulation of credits) and social integration (the attendance and involvement) are used to predict the persistence of the students. Meanwhile, the low importance of socioeconomic characteristics confirms the findings of Bean and Metzner (1985), according to which the environmental factors tend to be undervalued in the available data about institutions.

Deep Learning Architectures and Predictive Performance (RQ2)

The levels of accuracy of the analyzed tasks were also very diverse ranging from Cheng and Lin (2026) whose results vary between 66.67 and 73 % on a difficult and skewed Taiwanese task to that of Sangeetha and Priya (2026) whose accuracy stood at 98.58 % based on the data involving only one benchmark. Eight studies report either AUC-ROC or an AUC-PR equivalent, but none report the confidence interval or standard error needed for a pooled estimate, leading to the omission of the planned meta-analysis and the expression of its findings in a narrative form instead. Under those constraints; however, hybrid architectures that use a deep feature extractor combined with a second-stage ensemble or gradient boosting classifier deliver more consistent and stronger findings than single-architecture networks tested with similar data, which reflects the assumption that the second-stage classifier makes it more resistant to class imbalance resulting from the occurrence of the dropout prediction. Where a traditional algorithm is compared directly against a deep or hybrid model within the same study, the deep or hybrid model wins in every case, though the margin is smallest where a well-tuned XGBoost comparator, as in Aljawazneh et al. (2026), performs close to the deep learning model it is set against. This fits LeCun et al.'s (2015) view that deep learning builds its own representation instead of using preset features that is an advantage where predictors interact unpredictably. Bettahi et al. (2025) echo this and note that modular pipelines let representation and classification work together well.

Application of Explainable AI Techniques (RQ3)

Only five of the 26 studies applied any post-hoc explainability technique and a sixth, Ibrahim et al. (2026), names SHAP and LIME as future work rather than applying either. The SHAP appeared in four of the five, LIME in only one, alongside SHAP in Unnati et al. (2026). Baranyi et al. (2020) use SHAP to rank the twelve most influential features and note that this ranking mostly, though not entirely, agrees with a separate permutation importance measure. Rajput et al. (2025) apply SHAP at the level of individual predictions, reporting a SHAP value of plus 0.28 toward dropout for one modeled student driven by debt status and parental education. An example the authors use to argue that case-level explanation supports an advisor's decision in a way a global ranking alone cannot. Hamal et al. (2025) combine SHAP with Grad-CAM and find the two techniques converge on similar conclusions. This proportion, five of 26 studies or roughly 19 percent, sits far below Alharbi et al.'s (2026) finding that SHAP was applied in 74 of 75 machine learning studies, which suggests that explainability has not kept pace with the move toward deep learning specifically in this field, rather than being neglected in dropout prediction generally. Duro et al. (2026) similarly flag interpretability and deployment as unresolved across machine learning and deep learning alike. Albugami et al. (2024) add that explainability reporting usually stays global, a gap only Rajput et al. (2025) partly close here. The application rate of 19% for XAI in this case as opposed to 98% in the machine learning review conducted by Alharbi et al. (2026), which corroborates the claim made by Adadi and Berrada (2018) that the more complex the model is, the lesser the explainability is. Deep neural networks do not have built-in methods for interpretability, which creates a practice of showing the results of the model instead of being open to inspection.

Methodological Quality, Kenyan and African Relevance, and Research Gaps (RQ4)

The single Kenyan study by Kendagor et al. (2023) applied a convolutional neural network to a multi-university dataset and reports accuracy, precision and recall, but applies no explainability technique, does not clearly report how class imbalance was handled and was rated high risk of bias on the PROBAST analysis domain. It demonstrates that deep learning dropout prediction is achievable with Kenyan data, without providing a strong or explainable foundation to build directly on. Hamal et al.'s (2025) work on Morocco reports the highest outcomes in terms of methodological quality across three out of the four PROBAST dimensions, which is why the language spoken in the country, as well as the organizational framework and data structure, is considered an impediment to transferring the results to Kenya. It should be noted that using the adapted GRADE guidelines for judging evidence basing on its certainty shows low level of certainty for RQ1 and RQ2 due to certain known biases and limited representation of Africa. Read as a whole, the body of evidence supports the claim that deep learning outperforms traditional machine learning on this task and that academic and behavioral predictors carry the most consistent weight, but it does not yet support strong conclusions about which architecture, which explainability method, or which set of predictors would perform best in a Kenyan university, since so little of the underlying evidence was generated in a comparable setting. This matters in Kenya where socioeconomic pressure and weak first-year support are documented dropout drivers (Ogawa, 2024; Stofile & Enwereji, 2025) yet neither appears as a predictor in any of the 26 studies, leaving that part of the picture unexplained. The scant evidence from Africa, which forms only two of the 26 studies, seems to mirror the broken institutional data infrastructure, the scarce computational means, and the conventions that prefer already established international publishing avenues (Nyawira et al., 2025). This is not simply a gap of representation; it is also the case that the factors leading to dropout in Kenya are totally foreign to every other collection of predictors apart from the African ones.

Set against the two reviews that motivated this study, the findings here confirm rather than contradict the wider machine learning literature on one point and depart from it sharply on another. Academic and behavioral variables carry predictive weight most consistently in both this review and Lopez-Muñoz et al.'s (2025) broader catalogue of risk factors, a pattern reached independently through two different bodies of evidence. On explainability, however, the contrast is stark: Alharbi et al. (2026) found SHAP applied in 98 percent of the machine learning studies they reviewed, against 19 percent of the deep learning studies found here. Taken together, these two figures indicate that a gradient boosted tree or random forest arrived already carrying explainability tools that researchers picked up quickly, while a deep neural network or LSTM does not carry the same tooling as readily, and studies applying one appear more occupied with proving the model works than with explaining why. This is consistent with the account of explainable AI as a response to a gap between predictive strength and interpretability that grows as models grow more complex (Adadi & Berrada, 2018) and the evidence gathered here suggests that gap remains widest exactly where deep architectures have replaced the simpler models explainability tooling was originally built around. Neither this review nor the two it compares against found African higher education well represented, which points to a compounded gap this thesis was designed to address rather than one already closed by prior work.

Conclusions and Recommendations

The evidence gathered here shows that hybrid deep learning architectures, particularly those pairing deep feature extractor with a second-stage ensemble or gradient boosting classifier, reported the strongest predictive performance where directly compared against single-architecture alternatives and that deep or hybrid models outperform traditional machine learning comparators in every study that reported such a comparison directly. Academic and institutional or behavioral predictors, prior grades, credit accumulation, platform engagement and submission timeliness, recur most consistently across studies reporting predictor-level detail, while socioeconomic and demographic predictors appear less often and less consistently. In situations in which explainability has been applied, the prevailing technique is that of SHAP whereby both SHAP and LIME indicate that the same academic and behavioral variables come through as the strongest predictors. However, the conclusions in this regard cannot be made with a high degree of certainty for there was a low degree of certainty in all four questions as a result of the risk of bias in both result and analysis, the small number of studies that applied any technique of explainability and the limited evidence pointing towards Kenya.

This review indicates that deep learning may be the technique of choice for universities; however, the only study that has been done in Kenya has been determined to be highly biased and cannot be directly utilized in practice. In terms of student support services, the studies reviewed showed consistently that behavioral and academic signals are highly important factors to be taken into consideration in designing an early warning system, while socioeconomic and demographic characteristics are less relevant. In terms of policymaking, they should rather invest into improving the national data infrastructure, developing accurate outcome measures and eliminating biases in class sizes than attempt to apply a model that has not been validated and is based on a totally different country's framework.

Future research should prioritize four directions. Studies conducted in Kenyan and wider African institutions applying deep learning to a precisely defined dropout outcome, address the most urgent gap this review found, since the current evidence for that question rests on a single study. Studies applying SHAP, LIME, or both to a deep learning dropout model in any setting remain too few to support strong generalization. Studies reporting a full set of performance metrics for every model variant tested, rather than only the best-performing one, would close the reporting gap identified during this review. Finally, studies that test whether an advisor or administrator presented with a SHAP or LIME explanation actually finds it useful for designing an intervention would answer a question this review found almost entirely unaddressed, whether explainability changes what practitioners do or only what they can see.

Deep learning offers Kenyan universities a genuinely more effective tool for identifying students at risk of dropping out than most institutions currently use, and explainability offers a way to use that tool in a form advisors and administrators can trust. The evidence supporting the first half of that claim is reasonably strong, if unevenly distributed around the world. The evidence supporting the second half is thin everywhere and thinnest in the Kenyan context where it matters most. That gap is a reason to pursue the primary study this review was designed to prepare the ground for and not a reason to set it aside.

Acknowledgements

The author acknowledges the supervising academic who served as second reviewer throughout the screening, extraction, and risk-of-bias assessment process described in this review.

References

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). IEEE Access, 6, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

Albugami, S., Almagrabi, H., & Wali, A. (2024). From data to decision: Machine learning and explainable AI in student dropout prediction. Journal of e-Learning and Higher Education, 2024, Article 246301. https://doi.org/10.5171/2024.246301

Alharbi, A., Janarthanan, M., & Midhunchakkaravarthy, D. (2026). SHAP-centric interpretability in educational machine learning: A PRISMA systematic review of SHAP and LIME. Letters in Biomathematics.

Aljawazneh, H., Alqirem, R., & Al-Adwan, N. (2026). A comparative study of machine learning and deep learning algorithms for student dropout prediction in higher education. Al-Zaytoonah University Journal of Business. https://doi.org/10.15849/zjjb.v2i1.16

Altabrawee, H., Ali, O. A. J., & Ajmi, S. Q. (2019). Predicting students' performance using machine learning techniques. Journal of University of Babylon for Pure and Applied Sciences, 27(1), 194–205.

Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. Journal of Educational Data Mining, 1(1), 3–17. https://doi.org/10.5281/zenodo.3554657

Baranyi, M., Nagy, M., & Molontay, R. (2020). Interpretable deep learning for university dropout prediction. In SIGITE '20: Proceedings of the 21st Annual Conference on Information Technology Education. https://doi.org/10.1145/3368308.3415382

Bean, J. P., & Metzner, B. S. (1985). A conceptual model of nontraditional undergraduate student attrition. Review of Educational Research, 55(4), 485–540 https://doi.org/10.3102/00346543055004485

Bettahi, A., Belouadha, F.-Z., & Harroud, H. (2025). A modular and explainable machine learning pipeline for student dropout prediction in higher education. Algorithms, 18(10), 662. https://doi.org/10.3390/a18100662

Cheng, Y.-H., & Lin, C.-E. (2026). Predicting university student dropout risk using deep learning and ensemble voting mechanism. Engineering Proceedings. https://doi.org/10.3390/engproc2025120066

Chung, J. Y., & Lee, S. (2019). Dropout early warning systems for high school students using machine learning. Children and Youth Services Review, 96, 346–353. https://doi.org/10.1016/j.childyouth.2018.11.030

Commission for University Education. (2024). University statistics 2024/2025 report. https://www.cue.or.ke

Coppo, E. C., Caetano, R. S., de Lima, L. M., & Krohling, R. A. (2022). Student dropout prediction using 1D CNN-LSTM with variational autoencoder oversampling. In 2022 IEEE Latin American Conference on Computational Intelligence (LA-CCI). https://doi.org/10.1109/LA-CCI54402.2022.9981340

Debray, T. P. A., Damen, J. A. A. G., Snell, K. I. E., Ensor, J., Hooft, L., Reitsma, J. B., & Moons, K. G. M. (2017). A guide to systematic review and meta-analysis of prediction model performance. BMJ, 356, i6460. https://doi.org/10.1136/bmj.i6460

Duro, B., Gomes, A., Correia, F. B., Borges, A. R., & Bernardino, J. (2026). Machine learning and deep learning for dropout prediction in higher education: A review. Computers, 15(3), 164.

Gunning, D. (2017). Explainable artificial intelligence (XAI). Defense Advanced Research Projects Agency (DARPA), 2(2), 1.

Guyatt, G. H., Oxman, A. D., Akl, E. A., Kunz, R., Vist, G., Brozek, J., Norris, S., Falck-Ytter, Y., Glasziou, P., deBeer, H., Jaeschke, R., Rind, D., Meerpohl, J., Dahm, P., & Schunemann, H. J. (2011). GRADE guidelines: 1. Introduction, GRADE evidence profiles and summary of findings tables. Journal of Clinical Epidemiology, 64(4), 383–394. https://doi.org/10.1016/j.jclinepi.2010.04.026

Hamal, O., El Koufi, N., Cengiz, K., & Milić, L. (2025). AI-based prediction of student dropout using an enhanced deep learning framework: DeepDropNet model. International Journal of Advances in Soft Computing & Its Applications. https://doi.org/10.15849/IJASCA.251130.06

Ibrahim, A. M., Abonazel, M. R., & Ahmed, A.-H. N. (2026). A proposed hybrid deep learning and ensemble learning model for predicting student dropout in education. Statistics, Optimization & Information Computing. https://doi.org/10.19139/soic-2310-5070-3733

Kendagor, V. J., Chemwa, G., & Cheruyoit, W. (2023). A convolutional neural network framework for predicting student drop out in universities. In 2023 IEEE AFRICON. https://doi.org/10.1109/AFRICON55910.2023.10293554

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539

Lopez-Muñoz, G. L., Gonzalez Serrano, C., & Sanchez-Ferreira, C. (2025). Student dropout prediction in higher education: A systematic review of machine learning methods and risk factors. Journal of Information Technology Education: Research, 25, 1. https://doi.org/10.28945/5776

Loritee, L. J., Kyalo, D., & Nyaegah, J. O. (2026). Tuition fee levels and their influence on university student participation: Evidence from the University of Nairobi, Kenya. Journal of Education, 6(4), 16–28.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30.

Marcolino, M. R., Porto, T. R., Primo, T. T., Targino, R., Ramos, V., Queiroga, E. M., Munoz, R., & Cechinel, C. (2025). Student dropout prediction through machine learning optimization: Insights from Moodle log data. Scientific Reports, 15, 9840. https://doi.org/10.1038/s41598-025-93918-1

Mduma, N., & Machuve, D. (2021). Machine learning model for predicting student dropout: A case of Tanzania, Kenya and Uganda. In 2021 IEEE AFRICON (pp. 1–6). https://doi.org/10.1109/AFRICON51333.2021.9570956

Moons, K. G. M., de Groot, J. A. H., Bouwmeester, W., Vergouwe, Y., Mallett, S., Altman, D. G., Reitsma, J. B., & Collins, G. S. (2014). Critical appraisal and data extraction for systematic reviews of prediction modelling studies: The CHARMS checklist. PLoS Medicine, 11(10), e1001744. https://doi.org/10.1371/journal.pmed.1001744

Nagy, M., & Molontay, R. (2024). Interpretable dropout prediction: Towards XAI-based personalized intervention. International Journal of Artificial Intelligence in Education, 34(2), 274–300. https://doi.org/10.1007/s40593-024-00412-w

Nyawira, L., Musau, O., Adem, A., & Jobunga, E. (2025). Predicting student attrition in Kenyan universities: A comparative analysis of machine learning algorithms. Journal of the Kenya National Commission for UNESCO, 5(2), 1–9.

Ogawa, M. (2024). Massification of tertiary education and its inequality in Kenya: A case study of top students from a rural day secondary school. International Journal of Educational Development, 110, 103137. https://doi.org/10.1016/j.ijedudev.2024.103137

Oreopoulos, P., & Petronijevic, U. (2013). Making college worth it: A review of the returns to higher education. The Future of Children, 23(1), 41–65. https://doi.org/10.1353/foc.2013.0001

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan, a web and mobile app for systematic reviews. Systematic Reviews, 5, 210. https://doi.org/10.1186/s13643-016-0384-4

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hrobjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71. https://doi.org/10.1136/bmj.n71

Rajput, V., Singh, A., & Gupta, S. C. (2025). Early detection of student dropout risk via hybrid deep learning and gradient boosting algorithms. In 2025 International Conference on Engineering, Technology & Management (ICETM). https://doi.org/10.1109/ICETM63734.2025.11051847

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should I trust you? Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1135–1144). https://doi.org/10.1145/2939672.2939778

Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10(3), e1355. https://doi.org/10.1002/widm.1355

Sangeetha, K., & Priya, N. S. (2026). Improved GRU-based temporal modeling for robust student dropout classification and risk prediction. In 2026 International Conference on Smart Electronic Devices and Intelligent Systems (ICSEDIS). https://doi.org/10.1109/ICSEDIS68157.2026.11518148

Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. SN Computer Science, 2(3), 160. https://doi.org/10.1007/s42979-021-00592-x

Sivasankar, C., Palani, H. K., Fathima, A., Gunasekaran, K., Agoramoorthy, M., & Vasanthi, G. (2026). TERGATE-Warn: Temporal graph attention for equitable educational early warning. In 2026 6th International Conference on Expert Clouds and Applications (ICOECA). https://doi.org/10.1109/ICOECA68095.2026.11485035

Stofile, R., & Enwereji, C. (2025). Investigating the impact of institutional support services on student retention in a university in the Eastern Cape Province of South Africa. http://dx.doi.org/10.48047/jett.2025.16.06.07

Tinto, V. (2006). Research and practice of student retention: What next? Journal of College Student Retention: Research, Theory & Practice, 8(1), 1–19. https://doi.org/10.2190/4YNU-4TMB-22DJ-AN4W

Tinto, V. (2017). Through the eyes of students. Journal of College Student Retention: Research, Theory & Practice, 19(3), 254–269.

Unnati, M., Nakum, Z., & R, P. S. (2026). Hybrid CNN-LSTM model for student dropout risk prediction with interpretable AI. In 2026 International Conference on Computer Networks and Inventive Communication Technologies (ICCNCT). https://doi.org/10.1109/ICCNCT68477.2026.11590425

Wolff, R. F., Moons, K. G. M., Riley, R. D., Whiting, P. F., Westwood, M., Collins, G. S., Reitsma, J. B., Kleijnen, J., & Mallett, S. (2019). PROBAST: A tool to assess the risk of bias and applicability of prediction model studies. Annals of Internal Medicine, 170(1), 51–58. https://doi.org/10.7326/M18-1376