Discover Artificial Intelligence

RESEARCH

Open Access

Prediction of delayed postgraduate graduation using machine learning in Ugandan higher education

Musoke Robert Lubelenga[1,2*], Florence Nameere Kivunike[2], Halimu Chongomweru[2] and Emmanuel Ahishakiye[3]

*Correspondence:
Musoke Robert Lubelenga
robertmusoke21@gmail.com
[1]Department of Computer Science, Kyambogo University, Kampala, Uganda
[2]Department of Information Technology, Makerere University, Kampala, Uganda
[3]Department of Data Science and Artificial Intelligence, Kyambogo University, Kampala, Uganda

Abstract
Delayed postgraduate graduation is a major challenge to higher education institutions in Uganda in terms of student progression, institutional planning and human capital development. Identifying students at risk of delayed completion early can inform timely academic interventions and improve post-graduate outcomes. This study developed and compared machine learning models for delayed postgraduate graduation prediction using institutional administrative records from Makerere University. A retrospective predictive modelling study was performed using anonymised data from 500 postgraduate students registered from 2017 to 2023. 29% of them graduated on time and 71% graduated late. We developed and evaluated a number of supervised machine learning classifiers such as Random Forest, XGBoost, K-Nearest Neighbours, Logistic Regression and a stacking ensemble model. The model performance was evaluated using accuracy, precision, recall, F1-score and ROC-AUC. All trained models were found to have high and similar predictive performance in terms of accuracy (between ~94% and 95%) and ROC-AUC (between 0.94 and 0.95). Logistic Regression gave the numerically highest ROC-AUC (0.950) although differences in classifier performance were not statistically meaningful. Employment status and age at admission were positively related to delayed graduation whereas higher undergraduate CGPA decreased the likelihood of delayed completion. The results indicate that interpretable machine learning models can match the performance of complex ensemble methods on structured educational datasets. Apart from predictive performance, the study offers context-specific evidence on the use of predictive analytics in postgraduate education in Uganda, which is often characterised by patterns of progression affected by research-intensive study requirements, employment responsibilities, and long thesis completion processes. The proposed framework can assist in early identification of at-risk postgraduate students and directing institutional interventions in order to improve graduation outcomes.

Keywords Machine learning, Graduation prediction, Postgraduate education, Educational analytics, Uganda, Delayed completion

1 Introduction

Globally, the demand for postgraduate education has increased as students try to improve their career prospects [1]. Postgraduate programs provide opportunities for academic advancement, problem-solving, and skill development [2, 3]. But many postgraduate students are facing a challenge of not graduating on time [4]. On-time graduation means that the students complete the program tasks within the given time frame of the University [5, 6]. However, many postgraduate students worldwide do not finish their studies in time and developed countries spend on average four years instead of 2.5 years on their degrees [7, 8]. In sub-Saharan Africa, PG students may take three to four years to complete two-year master's degrees and a three-year PhD may take nine years [1, 9]. For example, research at the Makerere University Business School in Uganda found out that 95% of MBA students failed to graduate on time in 2015, 2016 and 2017 [9].

Postgraduate students do not complete their studies on time due to problems with coursework, time management, research skills, finances, changes in personal circumstances, motivation and skills, age, marital status, gender issues, health issues, lack of support, ineffective and untimely university intervention procedures, ineffective supervision and change in research focus [8, 10–13]. The non-completion of postgraduate students on time has serious implications for students, universities, and the wider education environment [14]. It makes students drop out at high rates and fail in their careers. It also causes loss of time and earnings [15–17]. Universities lose revenue, have lower graduation rates, inefficient allocation of resources, lower research productivity, and poor rankings [18]. The wider education landscape is facing a skills gap and reduced return on investment [19].

To solve the problem of non-graduation of postgraduate students on time, Ugandan universities adopted strategies including: increasing research supervisors, conducting seminars, providing grants and scholarships, creating research sub-committees, and providing guidelines to students [7, 9, 14]. Moreover, the government has also established institutions such as the National Council for Higher Education (NCHE) with its strategies to formulate guidelines to improve, among others, postgraduate graduation time [20]. Though this has been put in place, postgraduate students are graduating on time in Ugandan universities [1]. Over the past 10 years, there has been an increasing number of studies on student performance prediction worldwide, including at-risk students [21]. The prediction of students at risk of academic failure has great importance for timely institutional interventions in higher education to improve student performance and efficiency [22].

Early prediction of how long students will study and when they will graduate enables the identification of areas for improvement in learning [23]. Machine learning is used increasingly in education for prediction of student performance [24]. Machine learning algorithms can analyse data from students to find patterns that are associated with academic success, such as graduating on time and study duration [21, 25]. These studies commonly use classification models to predict student performance, with an emphasis on predicting at-risk students [26, 27]. Researchers are increasingly using machine learning to improve student outcomes and the effectiveness of higher education institutions [25, 28].

Ahmad and Shahzadi [29], proposed a machine learning-based model to find the answer to the question of whether the students were at risk regarding their academic performance. Rebai, et al. [30], proposed a machine learning-based model to identify the key factors affecting the academic performance of students and to determine the relationship between these factors. Yağcı [31], proposed a machine learning-based model to predict the final exam grades of students, taking their midterm exam grades as the source data. These studies show the usefulness of machine learning in educational prediction tasks, but most of the existing research mainly focuses on undergraduate education in developed or highly digitised educational environments. Postgraduate education in Uganda is marked by unique challenges, including the demands of research-intensive study, competing work responsibilities, limited funding opportunities, supervisory workload constraints, and lengthy thesis completion processes. These factors may have important implications for progression patterns and delays in completion that differ from undergraduate educational settings that are commonly represented in the existing educational data mining literature. Thus, there is limited evidence on the applicability and performance of machine learning methods for the prediction of delayed postgraduate graduation in Ugandan higher education institutions.

Thus, many researchers such as Albreiki et al. [25], Sekeroglu et al. [28], Alshanqiti and Namoun [32], claim that machine learning is potentially beneficial in predicting students' graduation time. We are aware of little or no research on these models for the prediction of graduation time for postgraduate students in Uganda. This study therefore exploits the predictive capacity of machine learning models to predict the graduation time of postgraduate students based on factors affecting the graduation time of students at Makerere University. ML algorithms have shown promise in predicting student performance, but further research is required to investigate their use in postgraduate programs [21].

This study develops and evaluates machine learning models for the prediction of delayed postgraduate graduation using institutional data from Makerere University. In addition, the study investigates whether complex ensemble methods provide any significant performance boost over simpler and more interpretable models in a structured educational dataset.

2 Related literature

2.1 Machine learning (ML) in education

One of the most promising applications of ML in education is the prediction of student learning outcomes [33]. By analyzing vast datasets of student performance data, demographics, and learning styles, ML models identifies patterns that correlate with success or potential difficulties. This predictive power allows identifying students at-risk of falling behind early on enables educators to provide targeted support, such as personalized learning plans or additional resources [27]. This proactive approach prevents small struggles from mounting into major learning gaps [28]. ML algorithms can personalize the learning experience by tailoring educational materials and assessments to individual student needs [21]. This caters to different learning styles and paces, fostering a more effective and engaging learning environment for all students. By identifying factors that contribute to student success, universities are able to optimize resource allocation [31, 34]. For example, additional support staff can be directed towards subjects with high prediction rates of student difficulty.

2.2 Prediction of student graduation-time using machine learning

Machine learning techniques have increasingly been applied to educational data mining to predict student progression, completion, and risk of delayed graduation. Most existing studies focus on undergraduate populations and primarily use supervised classification algorithms applied to institutional academic records.

Several researchers have applied individual classifiers such as Support Vector Machines, Naïve Bayes, and decision tree–based algorithms to predict student completion outcomes Bakri, et al. [35]; Lagman, et al. [36]; Bolsoni-Silva, et al. [37]; Wati and Indrawan [38]. These studies typically used demographic and academic performance indicators and reported moderate predictive performance, often within accuracy ranges of approximately 78%–92%. The findings consistently demonstrate that academic history and enrollment characteristics contain sufficient information to identify students at risk of delayed completion.

Other work has explored instance-based and modified classification methods, including reduced training vector support vector machines and multi-level classification frameworks for identifying at-risk students [39, 40]. While these approaches improve classification performance, they may require careful parameter tuning and can be sensitive to dataset size and class imbalance. More recent research has investigated ensemble learning approaches. Stacking and combined classifier models have been applied to student performance prediction using multiple base learners such as decision trees, k-nearest neighbors, Naïve Bayes, and support vector machines Saluja, et al. [41]. Ensemble methods aim to improve predictive stability by combining models with different learning biases, though they may increase computational complexity and are not always guaranteed to outperform well-regularized single classifiers.

Despite these developments, studies specifically examining postgraduate students remain limited. Postgraduate education differs substantially from undergraduate education in terms of student demographics, external commitments, and research-based programme structures. As a result, predictive models developed for undergraduate datasets may not generalize effectively to postgraduate populations. Furthermore, relatively few studies have examined structured administrative postgraduate records in low-resource higher education contexts. There is therefore a need to evaluate whether machine learning models can reliably identify students at risk of delayed postgraduate completion and to determine which factors are most strongly associated with timely graduation. Accordingly, this study applies a machine learning framework to postgraduate administrative records in order to (i) predict delayed graduation and (ii) identify the key predictors associated with completion-time.

2.3 Research gap and aim

Despite increasing application of machine learning in educational analytics, much of the existing literature focuses on improving predictive accuracy without evaluating whether complex models provide meaningful advantages over simpler and more interpretable approaches. In institutional higher-education datasets, which are typically small, structured, and administratively collected, model complexity may not necessarily translate into improved generalization performance. However, there is limited empirical evidence examining whether ensemble learning methods offer statistically significant predictive benefits over traditional models such as Logistic Regression in postgraduate education contexts, particularly in low-resource university settings. Therefore, rather than only developing a prediction system, this study investigates the comparative predictive value of complex ensemble learning relative to interpretable models for identifying postgraduate students at risk of delayed graduation. Using institutional academic records from Makerere University, the study evaluates whether stacking-based ensemble learning provides a measurable and statistically meaningful performance improvement over individual models. By doing so, the research contributes methodological insight into the suitability of advanced machine-learning techniques for small administrative educational datasets and informs the selection of transparent decision-support tools in higher-education management.

3 Materials and methods

3.1 Data source, variables, and sample selection

This study used retrospective postgraduate student administrative records obtained from Makerere University, Uganda, specifically from the College of Computing and Information Sciences, covering admissions between 2017 and 2023. The dataset was extracted from institutional academic registry databases and included demographic and academic attributes such as gender, age at admission, marital status, undergraduate academic performance (CGPA), year of undergraduate completion, admission year, employment status, funding type, study plan, and programme category. The original dataset contained minor missing values across selected variables, which are typical of administrative records. Rather than excluding incomplete records, missing values were addressed using imputation techniques during data preprocessing to preserve sample size and reduce potential bias.

Following data cleaning and preprocessing, the final analytical dataset consisted of 500 postgraduate student records, of which 145 students completed their programmes within the prescribed duration (on-time), and 355 experienced delayed completion. The outcome variable, Graduation Status, was defined as a binary classification indicating whether a student completed their programme within the official university completion-timeline (on-time) or exceeded the prescribed duration (delayed). For master's programmes, the official completion duration was two years, while students exceeding three years without completion were categorized as delayed in accordance with institutional postgraduate regulations. The modelling task was therefore formulated as a supervised classification problem. Predictor variables were restricted to demographic and academic attributes available at admission or during the early stages of study, allowing the model to function as an early-warning tool for identifying students at risk of delayed completion.

To prevent information leakage, outcome-derived variables and post-completion attributes such as time to completion, graduation year, completion duration, thesis submission date, and final academic progression records were excluded from model training. In addition, only students with sufficient follow-up duration relative to the approved programme completion-timeline were retained in the final analytical dataset. Consequently, students admitted during later cohorts (2022–2023) who had not yet exceeded the official completion duration and remained actively enrolled at the time of data extraction were excluded from delayed classification to minimize right-censoring bias and avoid misclassification of ongoing students as delayed completers.

The study population represents a subset of postgraduate students enrolled within the college during the study period. Makerere University has an estimated total enrolment of approximately 35,000–40,000 students, with postgraduate students constituting about 5–10% of the population. The College of Computing and Information Sciences reported an enrollment of approximately 2,386 students in the 2023/2024 academic year. The dataset used in this study, therefore, reflects a portion of the available postgraduate population within the college. Because the dataset was derived from institutional administrative records rather than a probabilistic sampling framework, it should be interpreted as a convenience sample of available records rather than a fully representative sample of all postgraduate students. Consequently, the findings provide insights into patterns of delayed graduation within the studied cohort but may be influenced by variations in record completeness and programme-level reporting practices.

All data were anonymized before access by the researchers. Personally identifiable information, including names, registration numbers, and contact details, was removed by the university administration. The study involved secondary analysis of de-identified institutional records, and no attempt was made to re-identify individuals. Institutional authorization for data use was granted in accordance with the university's data protection and confidentiality policies.

3.1.1 Predictor variables, outcome definition, and handling of right-censoring

The predictor variables used for model development were restricted to demographic, academic, and enrollment-related attributes available at admission or during the early stages of postgraduate study. These variables included age at admission, gender, marital status, undergraduate CGPA, undergraduate completion year, postgraduate admission year, employment status during study, funding type, study plan, and programme category. Table 1 summarizes the predictors included in the modelling framework. Importantly, variables directly derived from the outcome or containing post-completion information were excluded from model training to prevent information leakage. Specifically, variables such as time to completion, graduation year, completion duration, thesis submission date, and final academic status progression records were not used as predictors because they inherently encode information related to graduation outcomes. The predictive framework was therefore designed to simulate an early-warning system using only information that would reasonably be available before delayed graduation occurs.

Table 1 Predictor variables used in model development

Variable	Type	Description	Used in model
Age at admission	Numerical	Student age at admission	Yes
Gender	Categorical	Male/female	Yes
Marital status	Categorical	Single/married/other	Yes
Undergraduate CGPA	Numerical	Prior academic performance	Yes
Undergraduate completion year	Numerical	Year of prior degree completion	Yes
Postgraduate admission year	Numerical	Year admitted to postgraduate study	Yes
Employment status	Categorical	Employment during study	Yes
Funding type	Categorical	Self-sponsored/Scholarship	Yes
Study plan	Categorical	Coursework + Dissertation/Project	Yes
Programme category	Categorical	Postgraduate programme type	Yes
Time to completion	Numerical	Completion duration in months	No
Graduation year	Numerical	Year of graduation	No
Thesis submission date	Date	Final submission timing	No

The outcome variable, Graduation Status, was defined based on whether a student completed the programme within the officially prescribed institutional completion-timeline. Students who exceeded the permissible completion duration were classified as delayed, while those who completed within the institutional timeframe were classified as on-time graduates. To address potential right-censoring, special consideration was given to students admitted during the later years of the study period, particularly 2022–2023 cohorts. Students who had not yet exceeded the maximum allowable completion duration at the time of data extraction and who remained actively enrolled were excluded from delayed classification to avoid mislabeling ongoing students as delayed completers. Only students with sufficient elapsed follow-up time relative to the official programme completion window were included in the final modelling dataset. This approach minimized temporal bias and ensured that outcome labels reflected completed or legitimately delayed progression outcomes rather than incomplete observation periods.

3.2 Exploratory data analysis and variables for study

An exploratory data analysis was conducted to understand the structure and characteristics of the postgraduate student dataset before building the model. All descriptive statistics were calculated on the raw dataset of 500 students without any preprocessing, encoding, normalisation or class balancing to make sure that the values are representative of real-world measurements. The dataset contained variables related to postgraduate progression, including demographic, academic and enrolment related variables. The main features of the dataset are summarised in Table 2 and include numerical and categorical variables. The findings show that the average age of the postgraduate students is about 29.8 years, pointing to a blend of early and mid-career individuals. The academic performance of undergraduate population was good with average CGPA of 3.42 on 5 point scale. Also, the data indicate that a large number of students took a long time to graduate, with 71% not graduating on time. Further, most of the students were self-sponsored and most were employed during their studies indicating external commitments

Table 2 Summary of dataset characteristics ($N = 500$)

Variable	Category / statistic	Value
Age at admission (years)	Mean ± SD	29.8 ± 4.6
	Range	22–45
Undergraduate CGPA (5-point scale)	Mean ± SD	3.42 ± 0.48
	Range	2.10–4.50
Time to completion (months)	Mean ± SD	36.0 ± 8.2
	Range	24–60
Gender	Male	62.4%
	Female	37.6%
Marital status	Single	53.6%
	Married	38.4%
	Other	8.0%
Study plan	Plan A (Coursework + Dissertation)	67.0%
	Plan B (Coursework + Project Report)	33.0%
Funding type	Self-Sponsored	67.6%
	Scholarship-Sponsored	32.4%
Employment status	Employed	59.4%
	Not Employed	40.6%
Graduation status	On-time	29.0%
	Delayed	71.0%

Table 3 Missing values in the dataset before data cleaning

Variable	Missing count (*n*)	Missing (%)
Age at admission	12	2.4%
Undergraduate CGPA	18	3.6%
Employment status	10	2.0%
Undergraduate completion year	8	1.6%
Postgraduate admission year	15	3.0%
Graduation status	0	0.0%

Table 4 Class distribution before and after SMOTE resampling

Dataset version	Graduation status	Count (*n*)	Percentage (%)
Original dataset	On-time	145	29.0%
	Delayed	355	71.0%
Balanced dataset (SMOTE)	On-time	360	50.0%
	Delayed	360	50.0%

that may affect academic progression. The descriptive statistics reveal heterogeneity in students' characteristics and serve as a basis for understanding the factors which may influence the postgraduate graduation outcomes. These insights provided guidance on the choice of variables used in the predictive modelling stage.

3.3 Data processing and preparation

Data preparation was performed to ensure the dataset's quality and fitness for predictive modelling. This involved structural validation, consistency checks and removal of duplicate records that occurred as a result of administrative updates. A small amount of missing values was present in the dataset for selected variables (Table 3). Missing numeric variables were replaced with median values and categorical variables with the most frequent category to preserve sample size and minimise potential bias. The dataset was prepared for machine learning by z-score normalisation of the numerical variables and label encoding of the categorical variables into numerical form. All preprocessing steps such as imputation, normalisation and encoding were fitted on the training dataset and then applied to the validation and test datasets to reduce information leakage. The dataset was class imbalanced, with delayed graduation cases as the majority class. To solve this problem, the Synthetic Minority Over-sampling Technique (SMOTE) was applied only to the training data after data splitting. The class distribution before and after SMOTE resampling is summarised in Table 4. The class distributions of the validation and test datasets were kept as they were to prevent any bias in the model evaluation. Finally, we compared the predictive performance of the models using the original imbalanced dataset and the SMOTE-balanced training dataset, which generally yielded consistent predictive performance, arguing that the performance observed was not solely due to synthetic oversampling. We chose SMOTE over class weighting because SMOTE creates synthetic minority samples to better represent the minority class rather than adjusting penalties for misclassification. This increases the sensitivity of the classifier to under-represented graduation outcomes. Table 5 provides the results of Variance Inflation Factors (VIF) as a measure of multicollinearity among the predictor variables. The VIF values of all predictors were below the commonly accepted thresholds and no evidence of problematic multicollinearity was found in the modelling framework. These

Table 5 Variance inflation factor (VIF) assessment of predictor variables

Predictor variable	VIF	Interpretation
Age at admission	1.42	No evidence of problematic multicollinearity
Undergraduate CGPA	1.31	No evidence of problematic multicollinearity
Undergraduate completion year	1.27	No evidence of problematic multicollinearity
Postgraduate admission year	1.35	No evidence of problematic multicollinearity
Employment status	1.18	No evidence of problematic multicollinearity
Funding type	1.14	No evidence of problematic multicollinearity

preprocessing steps ensured consistency of the dataset, proper balance for model training and suitability for robust predictive analysis.

3.4 Predictive modelling approach

The final modelling framework included 10 predictor variables. The original dataset had 145 on-time graduation events which meant that an events-to-variables ratio was above the commonly recommended minimum ratios for stable logistic regression modelling. This ensured that the dataset was suitable for supervised classification analysis. In this paper, we developed a stacking ensemble learning framework to predict the graduation status (on-time or delayed) of postgraduate students using structured institutional academic and administrative data. Figure 1 illustrates the overall modelling workflow. The dataset was split into training (70%), validation (10%) and testing (20%) subsets following stratified sampling, which preserves the distribution of the outcome variable for all partitions, after data pre-processing. The Synthetic Minority Over-sampling Technique (SMOTE) was used to solve the class imbalance problem and this was applied only on the training dataset. The validation and test sets were kept in their original distributions to prevent data leakage. The modelling approach used 4 base classifiers, which are representative of different learning paradigms, namely Random Forest, XGBoost, K-Nearest Neighbours (KNN) and Logistic Regression. The models used are selected to represent the diversity of predictive patterns within the data set. Then a stacking strategy was used, whereby the predictions of the base models were aggregated and fed into a meta-learner. To achieve robust model training and prevent information leakage, we employed stratified 5-fold cross-validation on the training dataset to generate out-of-fold predictions for the meta-learner. We use Logistic Regression as a stable and interpretable meta-learner to combine the outputs of the models. The final ensemble model was evaluated on an independent test data set that was not used for training, resampling, or model selection. Standard classification metrics such as accuracy, precision, recall, F1-score and the area under the receiver operating characteristic curve (ROC-AUC) were used to assess performance. The model output was a binary classification of whether a student was predicted to graduate on-time or delayed completion.

3.5 Details of implementation and software environment

All analyses were conducted in Python (version 3.13). Data processing and analysis were performed with pandas and NumPy libraries. We used scikit-learn for machine learning models and gradient boosting was done using XGBoost. The stacking ensemble was implemented using the stacking classifier of scikit-learn. Class imbalance was addressed by using the Synthetic Minority Over-sampling Technique (SMOTE) as explained in

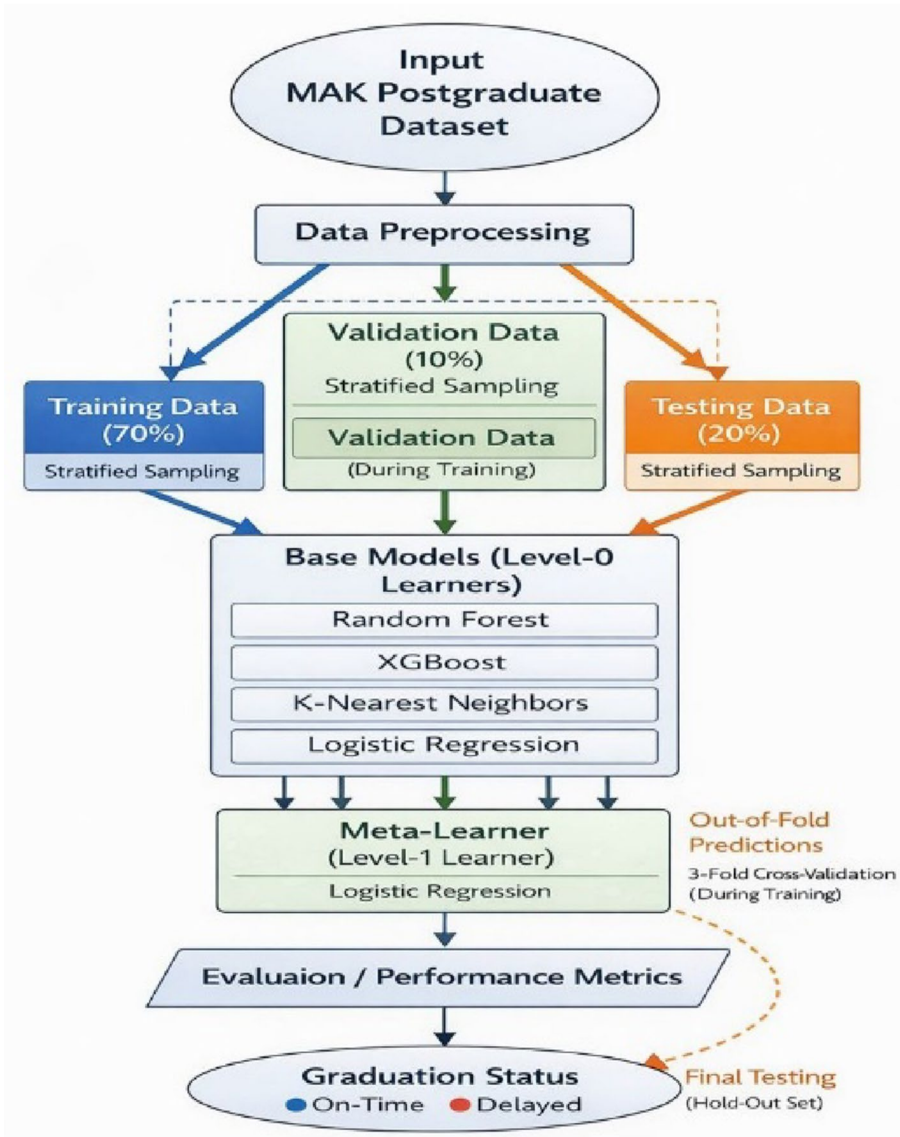

Fig. 1 The developed ensemble model

Table 6 Performance metrics for the developed ensemble classification model

Class	Accuracy	Precision	Recall	F1-Score
0 (On-time graduation)	0.948	0.948	0.950	0.949
1 (Delayed graduation)	0.948	0.948	0.946	0.947
Average	0.948	0.948	0.948	0.948

Sect. 3.4.4. For reproducibility, all models were trained and tested on the same workstation under the same computational conditions (Intel Core i7, 16 GB RAM).

4 Results

4.1 Stacking ensemble model result

The stacking ensemble model was tested on the separate held-out test data set that contained only original student records not utilised during model training. Accuracy, precision, recall and F1-score were used to evaluate the performance. Table 6 summarises the classification results. The model demonstrated balanced classification performance between the delayed and on-time completion categories, with similar precision and recall values for the two classes. This indicates that the classifier was able to detect delayed students without greatly increasing misclassification of on-time students. The overall accuracy was approximately 94.8% indicating good discrimination between the completion outcomes. The closeness of macro-averaged and weighted-averaged metrics indicates that the model did not over-favor any of the classes, despite the original dataset imbalance. The results confirm that the model learned consistent decision boundaries, as opposed to majority-class prediction. However, understanding the performance of the ensemble model necessitates comparison with individual classifiers in the subsequent sections to see if combining multiple models resulted in any meaningful predictive advantage.

4.2 Comparison of classification model performance

The predictive performances of the machine-learning classifiers were compared on the independent held-out test dataset. The performance parameters that were tested were accuracy, precision, recall, F1-score and ROC-AUC. The results of the comparison are summarised in Table 7. All the evaluated models showed consistent high predictive performance, with accuracy values between approximately 94% and 95%. The numerically highest ROC-AUC was obtained by Logistic Regression. Random Forest, XGBoost, K-Nearest Neighbours and the stacking ensemble performed similar in all evaluation metrics. The observed differences between the models were small and the 95% confidence intervals overlapped, suggesting no statistically meaningful differences in predictive performance. The results indicate that the predictive performance was mainly driven by the information included in student attributes and not by the increase in the algorithmic complexity. Although the Logistic Regression performed better in terms of ROC-AUC, the stacking ensemble did not present a significant predictive advantage over simpler and more interpretable models. All reported values were calculated on an independent held-out test set and 95% confidence intervals were estimated using bootstrap resampling.

To further evaluate the influence of class imbalance handling, a comparison was performed on the original imbalanced dataset and the SMOTE-balanced training dataset. Table 8 shows the summary of the comparative results. SMOTE increased classification sensitivity and overall predictive balance slightly, but did not significantly change the

Table 7 Comparative Performance of Evaluated Classification Models with 95% Confidence Intervals

Model	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F1-score (95% CI)	ROC–AUC (95% CI)
Logistic regression	0.948 (0.921–0.972)	0.948 (0.920–0.971)	0.947 (0.919–0.970)	0.947 (0.919–0.970)	0.950 (0.928–0.971)
Random forest	0.948 (0.920–0.971)	0.947 (0.919–0.970)	0.947 (0.919–0.970)	0.947 (0.919–0.970)	0.946 (0.923–0.968)
XGBoost	0.947 (0.919–0.970)	0.947 (0.918–0.969)	0.946 (0.917–0.968)	0.946 (0.917–0.968)	0.945 (0.921–0.967)
K-Nearest neighbors	0.946 (0.917–0.969)	0.946 (0.917–0.968)	0.945 (0.916–0.967)	0.945 (0.916–0.967)	0.944 (0.920–0.966)
Stacking ensemble	0.948 (0.921–0.972)	0.948 (0.920–0.971)	0.948 (0.920–0.971)	0.948 (0.920–0.971)	0.947 (0.924–0.969)

Table 8 Comparative model performance with and without SMOTE

Model	Accuracy without SMOTE	Accuracy with SMOTE	ROC-AUC without SMOTE	ROC-AUC with SMOTE
Logistic regression	0.941	0.948	0.943	0.950
Random forest	0.940	0.948	0.941	0.946
XGBoost	0.939	0.947	0.940	0.945
K-nearest neighbors	0.938	0.946	0.939	0.944
Stacking ensemble	0.942	0.948	0.944	0.947

Table 9 Cross-validation performance comparison of evaluated classification models

Model	Accuracy (Mean ± SD)	Precision (Mean ± SD)	Recall (Mean ± SD)	F1-score (Mean ± SD)
Logistic regression	0.948 ± 0.002	0.948 ± 0.002	0.947 ± 0.002	0.947 ± 0.002
Random forest	0.947 ± 0.003	0.947 ± 0.003	0.946 ± 0.003	0.946 ± 0.003
XGBoost	0.946 ± 0.003	0.946 ± 0.003	0.945 ± 0.003	0.945 ± 0.003
K-nearest neighbors	0.945 ± 0.004	0.945 ± 0.004	0.944 ± 0.004	0.944 ± 0.004
Stacking ensemble	0.948 ± 0.003	0.948 ± 0.003	0.947 ± 0.003	0.947 ± 0.003

relative performance patterns of classifiers, indicating that the observed predictive performance was not solely due to synthetic oversampling.

4.2.1 Comparison of statistical results and model stability analysis

The differences in the predictive performance between the evaluated classifiers were small, especially for the ROC-AUC metric, the values of which ranged between approximately 0.94 and 0.95. The overlapping 95% confidence intervals across models suggested no statistically meaningful differences in predictive discrimination performance. In the development of the model, stratified k-fold cross-validation was used to assess model stability and reduce dependence on a particular train-test split. In each iteration, one fold was used for validation and the remaining folds were used for training, and performance metrics were averaged across folds. The results of the cross-validation are summarised in Table 9. All models were evaluated and showed a very consistent performance across folds with low standard deviations for accuracy and F1 score. In addition, the results of the independent test were consistent with the cross-validation estimates, indicating stable generalisation performance and less risk of overfitting or information leakage. Results also indicate that the predictive performance was driven mainly by the information contained in the student attributes and not by increased algorithmic complexity. Values in Table 6 are the mean and standard deviation from the stratified cross-validation on the training dataset.

4.3 Confusion matrix analysis

Figure 2 shows the pooled out-of-fold confusion matrices using the stratified five-fold cross validation on all the 500 student records. For this evaluation, a student was only classified if the student belonged to a held-out fold, i.e. the model corresponding to this fold was not trained on this record. The sample consisted of 145 on-time graduates and 355 delayed graduates. The Random Forest correctly classified 137 students in time and 337 students delayed, with 26 misclassifications and an accuracy of 94.8%. The XGBoost model correctly identified 138 on-time students and 336 delayed students, with an overall accuracy of 94.8%. K-Nearest Neighbours performed slightly worse correctly

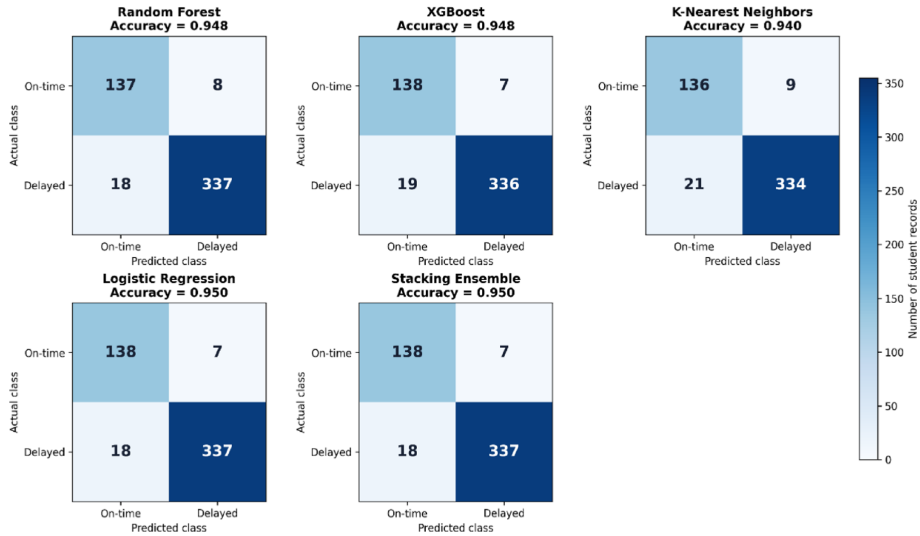

Fig. 2 Confusion matrix of the stacking ensemble model and base models

classifying 136 on-time students and 334 delayed students with 30 misclassifications and an accuracy of 94.0%. The stacking ensemble had the same classification results as the Logistic Regression. Both models correctly classified 138 out of 145 on time students and 337 out of 355 delayed students. This translates into 25 misclassifications and an accuracy of 95.0%. However, the stacking ensemble did not reduce the classification errors compared to Logistic Regression, although it obtained a strong and balanced classification performance. This suggests that for this dataset, the ensemble did not provide any substantial practical advantage over the simpler and more interpretable Logistic Regression model.

4.4 Receiver operating characteristic (ROC) curve analysis

Figure 3 presents the Receiver Operating Characteristic curves comparing the discrimination ability of the stacking ensemble and the individual classifiers. All the models produced curves close to the top left corner of the plot, indicating high sensitivity and low false positive rates. Area under the receiver operating characteristic curve (ROC-AUC) confirmed robust and consistent performance for all models. The Logistic Regression model achieved the highest ROC-AUC score of 0.9467, indicating a slightly better ability to differentiate between students who graduate on-time and those who have delays. Likewise, Random Forest, XGBoost and K-Nearest Neighbours, had ROC-AUC values near 0.94, which is a high predictive capability. The stacking ensemble achieved a ROC-AUC score of 0.9419. This is a good classification performance, but does not outperform the best individual classifier. All ROC curves are very close to each other, suggesting that the complexity of the models did not have a significant impact on the discrimination performance. This suggests that less complex models can perform as well as complex ensembles over structured institutional administration datasets.

4.5 Importance of features

We conducted a feature importance analysis using the Random Forest classifier and impurity based importance scores to identify variables that contributed most strongly to model predictions. As can be seen in Fig. 4, employment status, age at admission and

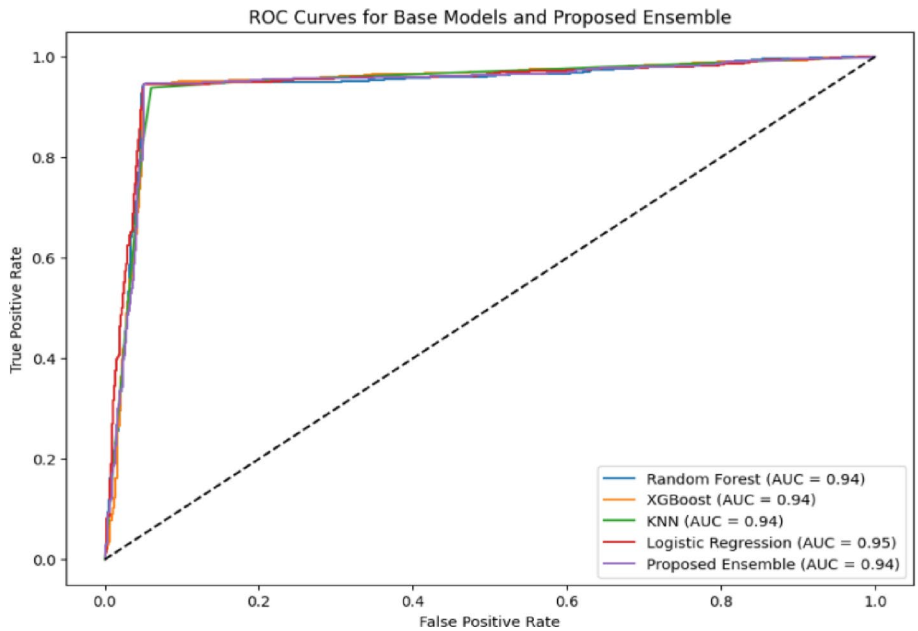

Fig. 3 The ROC curve

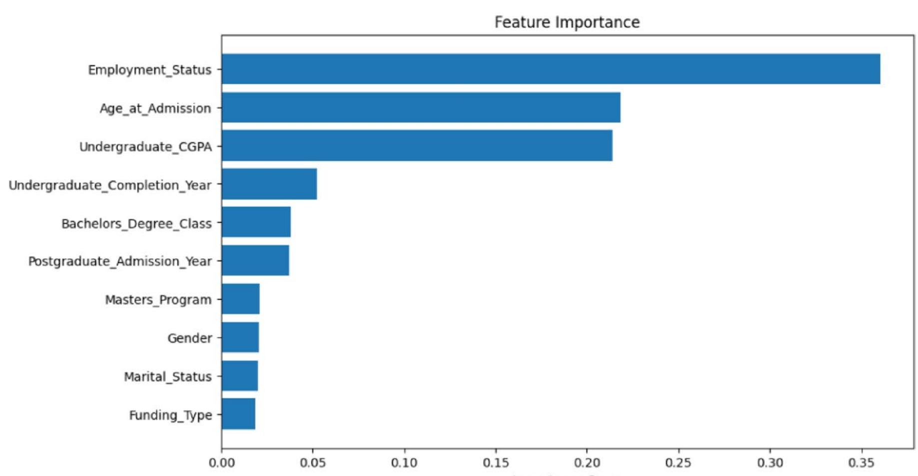

Fig. 4 Random forest feature importance rankings for predictor variables

undergraduate CGPA were the variables most strongly related to the model's classification output. Note that feature importance refers to predictive association, not causal influence. Variables identified should be viewed as predictors of increased or decreased likelihood of delayed completion, rather than as direct causes of delayed graduation. For example, the employment status may be a reflection of competing time commitments, while undergraduate CGPA may be a reflection of prior academic preparedness. Lower ranked variables such as gender, marital status and funding type contributed relatively less to classification performance, suggesting limited additional predictive information beyond the core academic and administrative variables.

4.6 Tuning of base learners' hyperparameters

The model stability and overfitting were reduced by tuning hyperparameters. The base classifiers (Random Forest, XGBoost, K-Nearest Neighbours and Logistic Regression) were optimised using grid search with stratified 5-fold cross-validation on the training dataset. The optimisation metric was set as macro-averaged F1-score to ensure that the classifier performed equally well on both graduation outcome classes. Random Forest was tested with a combination of n_estimators = [50, 100, 200], max_depth = [5, 10, 15], min_samples_split = [2, 5, 10]. The best configuration is 100 trees with a maximum depth of 10. XGBoost was optimised for n_estimators = [50, 100, 200], max_depth = [3, 6, 9], learning_rate = [0.01, 0.1, 0.2], subsample = [0.8, 1.0], colsample_bytree = [0.8, 1.0]. The best configuration was 100 trees, maximum depth 6, and learning rate 0.1 with feature and row sub-sampling. For the K-Nearest Neighbours classifier the search space was n_neighbors = [3, 5, 7, 9] and distance metrics = [Euclidean, Manhattan]. The best configuration was obtained with 7 neighbours. Logistic Regression: Tuned with regularisation strengths C = [0.01, 0.1, 1, 10] and solvers = [lbfgs, liblinear] The best result was obtained with C = 0.1 and lbfgs solver. The optimised models were retrained and added to the stacking ensemble after tuning as summarised in Table 10. Tuning improved the predictive stability across classifiers, but the performance differences between ensemble and individual models remained small. Logistic Regression continued to provide the numerically highest ROC-AUC values, indicating that the predictive performance was primarily driven by the information contained in the student attributes, rather than increased algorithmic complexity.

4.7 Feature impact direction (interpretability analysis)

For interpretability, the coefficients from the Logistic Regression were reported as odds ratios with the corresponding 95% confidence intervals. Odds ratios greater than 1 indicate an increased likelihood of delayed graduation, whereas odds ratios less than 1 indicate a decreased likelihood of delayed graduation. In Table 11, we present the associations of the main predictors with delayed graduation outcomes. Findings suggest that working while studying was associated with a higher probability of delayed completion

Table 10 Optimized hyperparameters for base learners in the stacking ensemble using grid search with stratified 5-fold cross-validation

Model	Tuning method	Evaluation metric	Hyperparameter search space	Optimized hyperparameters
Random forest	Grid search + stratified 5-Fold CV	F1-macro	n_estimators = [50, 100, 200]; max_depth = [5, 10, 15]; min_samples_split = [2, 5, 10]	n_estimators = 100; max_depth = 10; min_samples_split = 2
XGBoost	Grid search + stratified 5-Fold CV	F1-macro	n_estimators = [50, 100, 200]; max_depth = [3, 6, 9]; learning_rate = [0.01, 0.1, 0.2]; subsample = [0.8, 1.0]; colsample_bytree = [0.8, 1.0]	n_estimators = 100; max_depth = 6; learning_rate = 0.1; subsample = 0.8; colsample_bytree = 0.8
K-nearest neighbors (KNN)	Grid search + stratified 5-Fold CV	F1-macro	n_neighbors = [3, 5, 7, 9]; distance metric = [Euclidean, Manhattan]	n_neighbors = 7
Logistic regression	Grid search + stratified 5-Fold CV	F1-macro	C = [0.01, 0.1, 1, 10]; solver = [lbfgs, liblinear]	C = 0.1; solver = lbfgs
Proposed stacking ensemble	Retraining with optimized base models	Accuracy, F1-score, ROC-AUC	Inputs from optimized base learners	Combined predictions from optimized base models

Table 11 Logistic regression odds ratios and association with delayed graduation

Predictor	Odds ratio (95% CI)	Interpretation
Employment status (employed)	2.14 (1.42–3.21)	Employed students were more likely to experience delayed graduation
Age at admission	1.36 (1.12–1.68)	Older students showed increased likelihood of delayed completion
Undergraduate CGPA	0.58 (0.41–0.82)	Higher prior academic performance was associated with reduced risk of delay
Funding (scholarship)	0.64 (0.45–0.91)	Sponsored students were more likely to complete on-time
Gender	1.05 (0.79–1.38)	Minimal association with delayed graduation was observed

potentially due to juggling competing work and academic demands. In addition, older age at entry was associated with delayed graduation. Conversely, a higher undergraduate CGPA was associated with a protective effect, such that students with better academic preparation had a higher likelihood of completing their programs on time. Scholarship funding was also associated with a lower likelihood of delayed completion. Note that these findings are to be interpreted as predictive associations, and not as causal effects. It is a model that looks for patterns that can be used to identify risk early, but does not establish that these variables cause delayed graduation.

5 Discussion of results

This study investigated the performance of several machine learning models to predict delayed postgraduate graduation based on institutional administrative data. The results showed that all models tested had high and comparable predictive performance, and more complex ensemble models did not significantly outperform simpler models. This result implies that algorithmic complexity is not the main driver of predictive performance for structured educational datasets but rather the quality and structure of the underlying data. In such contexts, interpretable models, e.g., Logistic Regression, might be sufficiently predictive while also being more transparent for academic decision-making. This is especially true in higher education contexts where stakeholders need clear and explainable insights to inform intervention strategies.

Consistent with prior work in educational data mining that shows model performance is more affected by dataset characteristics, feature selection, and preprocessing than by the algorithm itself [42–44]. Similarly, applied machine learning literature has demonstrated that for tabular data, simple models can achieve performance comparable to complex ensemble methods, provided that the available features encode predictive signals of interest [43, 45, 46].

The strong performance of Logistic Regression in this study confirms its suitability for structured prediction tasks. Even as more complex algorithms become more popular, linear models can still do a good job, if the relationships between predictors and outcomes are stable and are captured well in the data. Moreover, the interpretability of Logistic Regression is a plus for academic decision-support systems that need transparency and ease of interpretation to be adopted by institutions.

The analysis also identified important predictors associated with delayed completion of postgraduate studies. Delayed graduation was positively associated with employment status and age at admission, suggesting external responsibilities such as work and family commitments may impact academic progression. In contrast, those who graduated on time had a higher undergraduate CGPA, suggesting prior academic preparation is an important determinant of postgraduate success. These findings are in line with earlier studies on the influence of work commitments and academic background on postgraduate progress [42–44], and studies on the influence of age and external responsibilities on completion outcomes [47, 48].

Postgraduate education in Uganda is different from undergraduate education in the research-intensive nature of study requirements, competing employment responsibilities, limited funding opportunities, supervisory workload constraints and extended thesis completion processes. Progression patterns are more complex and less predictable than in traditional undergraduate contexts, with many postgraduate students studying while in full-time employment. The value of this study therefore lies not so much in the novelty of the machine learning algorithms used but in the contextual use of predictive analytics in the relatively under-researched field of postgraduate education in Uganda. The findings therefore add context-specific evidence of the use of institutional administrative data to identify early postgraduate students at risk for delayed completion in resource-constrained higher education environments.

On the other hand, demographic and programme-related variables such as gender, marital status and funding type contributed little to the prediction performance. This suggests that demographic characteristics may have less impact on completion time for the current study than academic preparedness and external commitments, which is consistent with mixed findings in the literature [43, 45, 46].

It is important to note that the relations found in this study are predictive associations and not causal effects. The models identify patterns in past data to help identify students who may be at risk for delayed completion, but do not suggest direct causal mechanisms. These variables should thus be interpreted as measures of early intervention rather than determinants of student outcome.

The results provide support for the use of data-informed monitoring systems in postgraduate education. Institutions can use administrative data that they routinely collect to identify students at risk early and to implement targeted support strategies, such as academic advising, supervision support and workload adjustments. This enables a shift from reactive to proactive management of academics.

The predictive framework can be implemented with a probability based risk threshold for practical deployment of identifying students in need of early intervention. In this study, a probability threshold of 0.50 can identify students at high risk of delayed graduation. Students who surpass this threshold could then be considered for targeted support interventions such as academic advising, supervision support, mentoring or workload management. Institutions can also adjust the threshold depending on the support capacity available and the level of sensitivity to at-risk cases they wish to achieve.

The consequences of classification errors in the course of institutional deployment should also be taken into account. False positives may identify students as at risk for delayed completion when they are not, leading to unnecessary intervention or administrative concern. But such interventions might still be a useful academic support in supportive educational settings. False negatives, however, could lead to the failure to identify students who are at risk, thereby preventing timely and appropriate intervention by the institution. As such, the predictive framework should be used as a decision-support tool that augments, rather than replaces, academic judgement and mechanisms to support students.

Responsible use of predictive analytics in higher education involves grappling with student privacy, transparency, fairness and data governance issues. Institutions adopting the proposed framework should use predictions only to support academic interventions and not for punitive decision making. In an institutional deployment, there should be human oversight, periodic model evaluation and mechanisms to monitor for potential bias across groups of students. Responsible implementation still means the use of anonymised administrative data and the application of institutional data protection policies.

However, the results should be considered given the limitations of observational educational data. The dataset is from one college of one university and may have academic, administrative and demographic features unique to that institution. Thus the developed models should be seen as proof-of-concept rather than a universally generalisable predictive framework. Furthermore, administrative data can still have underlying data quality limitations even after the preprocessing and imputation methods applied in this study. In addition, although SMOTE can improve the class balance during model training, the synthetic oversampling may lead to localised patterns that are not representative of the underlying population distribution.

Future research should explore the proposed framework in larger multi-institutional and prospective datasets to assess its generalisability across broader higher education contexts. Validation of the model and generalisability should be the focus of future work using prospective cohort and multi-institutional data sets. Adding longitudinal data on progression in academia could improve predictive accuracy even further and allow continuous tracking of students' performance over time.

6 Conclusion

The study developed and evaluated machine learning models for predicting delayed graduation of postgraduates using institutional administrative data at Makerere University. All models under evaluation performed well and similarly in terms of predictive performances, with accuracy values comprised between approximately 94 and 95% and ROC-AUC values comprised between 0.94 and 0.95. Logistic Regression achieved the numerically highest ROC-AUC of 0.950. Differences between classifiers were not statistically meaningful. The results also identified employment status, age at admission and undergraduate CGPA as the most significant predictors associated with delayed postgraduate completion. These results show that regularly gathered institutional data can assist in the early identification of postgraduate students at risk of delayed completion and guide timely institutional interventions. The study provides context-specific evidence on the application of predictive analytics in postgraduate education in Uganda. The progression patterns might be influenced by the requirements of research-intensive programmes, employment responsibilities, funding limitations and long processes of thesis completion.

However, the results need to be interpreted in the context of the limitations of this study. Limitations and Future Directions The dataset was collected from a single college in a single university and may not generalise to other institutional settings. Furthermore, the use of retrospective observational data restricts interpretation to predictive associations rather than causal relationships. Also, while SMOTE improved class balance during training, synthetic oversampling may create local patterns that do not fully capture the underlying population distribution.

Future studies should aim at validating the proposed framework in larger, multi-institutional and prospective data and include longitudinal data about the academic progression to facilitate continuous monitoring and improve the model generalisability and robustness.

Acknowledgements
None.

Author contributions
The authors confirm their contribution to the paper as follows: study conception and design: Musoke Robert Lubelenga (MRL); data collection: MRL; analysis and interpretation of results: MRL, Emmanuel Ahishakiye(EA); draft manuscript preparation: MRL; Supervision: Florence Nameere Kivunike (FNK), Halimu Chongomweru (HC); Manuscript review and editing: MRL, EA. All authors reviewed the results and approved the final version of the manuscript.

Funding
None.

Data availability
The dataset analyzed during the current study contains confidential institutional student records and cannot be publicly shared. Access may be granted by the university administration subject to institutional approval and data protection regulations.

Code availability
Not applicable.

Declarations

Ethics approval and consent to participate
The need for ethical approval and informed consent was waived by the Makerere University School of Computing and Informatics Technology due to the retrospective nature of the study and the use of fully anonymized institutional administrative data. The study complied with institutional data protection and confidentiality guidelines.

Consent for publication
Not applicable.

Use of artificial intelligence
The authors used generative artificial intelligence (ChatGPT) to assist with language refinement and manuscript editing. The authors take full responsibility for the accuracy and integrity of the content.

Competing interests
The authors declare no competing interests.

Received: 12 January 2026 / Accepted: 7 August 2026
Published online: 22 August 2026

References

1. Ssenyonga J, Nakiganda PB. Postgraduate student research realities in Uganda. Postgraduate Research Engagement in Low Resource Settings. ed: IGI Global; 2020. pp. 150–72.
2. Atibuni DZ, Kibanja GM, Olema DK, Ssenyonga J, Karl S. Challenges and strategies of research engagement among Master of Education students in Uganda, 2017.
3. Roach M, Sauermann H. The declining interest in an academic career. PLoS ONE. 2017;12:e0184130.
4. Ismail AM. A qualitative analysis of factors affecting the successful completion of postgraduate studies. Turkish J Comput Math Educ (TURCOMAT). 2021;12:152–61.
5. Aparicio M, Oliveira T, Bacao F, Painho M. Gamification: a key determinant of massive open online course (MOOC) success. Inf Manag. 2019;56:39–54.
6. Burns T, Sinfield S. Essential study skills: the complete guide to success at university, 2022.
7. Malunda PN, Atwebembeire J, Ssentamu PN. Research supervision as an antecedent to graduate student progression in the public higher institutions of learning in Uganda. Int J Learn Teach Educational Res, 20, 2021.
8. Namubiru P, Japheth N. Invest in research supervision, enhance timely completion of postgraduate studies. RMC J Social Sci Humanit. 2021;2:0403.
9. Waweru IK, Kyakuha M. Attitude towards research among masters students in Makerere University Business School (MUBS), Uganda. Int J Mod Educ Stud. 2020;4:42–56.
10. Shahsavar Z, Kourepaz H. Postgraduate students' difficulties in writing their theses literature review. Cogent Educ. 2020;7:1784620.
11. Nwachukwu PT, Thobejane TD, Nyathi NA, Iwara I, Obadire O, Kilonzo B et al. Author pages title of articles, 2018.
12. Desmennu AT, Owoaje E. Challenges of research conduct among postgraduate research students in an African University. Educational Res Rev. 2018;13:336–42.
13. Hoon TS, Narayanan G, Sidhu GK, Choo LP, Fook CY, Salleh N. Students' perceptions toward postgraduate study: a preliminary investigation. Int J Educ. 2019;4:123–38.

14. Kayondo. Effect of monitoring and evaluation processes on student course completion in Universities. Int J Technol Manage. 2020;5:15–15.
15. Carnevale AP, Smith N. Balancing work and learning: implications for low-income students, 2018.
16. Muthukrishnan P, Gurnam K, Hoon T, Geethanjali N, Chan YF. Key factors influencing graduation on time among postgraduate students: a PLS-SEM Approach. Asian J Univ Educ. 2022;18:0214.
17. Priyadarshini M, Gurnam KS, Hoon TS, Geethanjali N, Fook CY. Key factors influencing graduation on time among postgraduate students: a PLS-SEM Approach. Asian J Univ Educ. 2022;18:51–64.
18. Casanova VS, Pullido ML. Factors of graduate students' attrition and retention in occidental Mindoro State College Graduate School. Int J Educational Res Social Sci (IJERSC). 2022;3:826–31.
19. Kaiser SK. Employer reports of skills gaps in the workforce. The University of Nebraska-Lincoln; 2019.
20. NationalCouncilforHigherEducation SP. 2020/2021–2024/2025, N. C. f. H. Education, Ed., ed, 2020.
21. Karalar H, Kapucu C, Gürüler H. Predicting students at risk of academic failure using ensemble model during pandemic in a distance learning system. Int J Educational Technol High Educ. 2021;18:63.
22. Yuliansyah H, Imaniati RAP, Wirasto A, Wibowo M. Predicting students graduate on time using C4. 5 algorithm. J Inf Syst Eng Bus Intell. 2021;7:67–73.
23. Zhang Y, Ghandour A, Shestak V. Using learning analytics to predict students performance in moodle LMS, 2020.
24. Mduma N, Machuve D. Machine learning model for predicting student dropout: a case of Tanzania, Kenya and Uganda, in 2021 IEEE AFRICON, 2021, pp. 1–6.
25. Albreiki B, Zaki N, Alashwal H. A systematic literature review of student'performance prediction using machine learning techniques. Educ Sci. 2021;11:552.
26. Sen PC, Hajra M, Ghosh M. Supervised classification algorithms in machine learning: a survey and review, In Emerging Technology in Modelling and Graphics: Proceedings of IEM Graph 2018, 2020, pp. 99–111.
27. Nawang H, Makhtar M, Hamzah W. A systematic literature review on student performance predictions. Int J Adv Technol Eng Explor. 2021;8:1441–53.
28. Sekeroglu B, Abiyev R, Ilhan A, Arslan M, Idoko JB. Systematic literature review on machine learning and student performance prediction: critical gaps and possible remedies. Appl Sci. 2021;11:10907.
29. Ahmad Z, Shahzadi E. Prediction of students' academic performance using artificial neural network. Bull Educ Res. 2018;40:157–64.
30. Rebai S, Yahia FB, Essid H. A graphically based machine learning approach to predict secondary schools performance in Tunisia. Socio-Economic Plann Sci. 2020;70:100724.
31. Yağcı M. Educational data mining: prediction of students' academic performance using machine learning algorithms. Smart Learn Environ, 9, 11, 2022.
32. Alshanqiti A, Namoun A. Predicting student performance and its influential factors using hybrid regression and multi-label classification. IEEE Access. 2020;8:203827–44.
33. Aiken JM, De Bin R, Hjorth-Jensen M, Caballero MD. Predicting time to graduation at a large enrollment American university. PLoS ONE. 2020;15:e0242334.
34. Trivedi S. Improving students' retention using machine learning: impacts and implications. ScienceOpen Preprints, 2022.
35. Bakri R, Astuti N, Ahmar A. Machine learning algorithms with parameter tuning to predict students' graduation-on-time: a case study in higher education. J Appl Sci Eng Technol Educ. 2022;4:259–65.
36. Lagman AC, Calleja JQ, Fernando CG, Gonzales JG, Legaspi JB, Ortega JHJC et al. ,., Embedding naïve Bayes algorithm data model in predicting student graduation, presented at the Proceedings of the 3rd International Conference on Telecommunications and Communication Engineering, Tokyo, Japan, 2020.
37. Bolsoni-Silva AT, Barbosa RM, Brandão AS, Loureiro SR. Prediction of course completion by students of a university in Brazil. Psico-USF. 2018;23:425–36.
38. Wati M, Indrawan W. Predicting degree-completion time with data mining, in 2017 3rd International Conference on Science in Information Technology (ICSITech), 2017, pp. 732–736.
39. Zhang S, Wang. Machine learning and AI in cancer prognosis, prediction, and treatment selection: a critical approach. J Multidiscip Healthc. 2023;16:1779–91.
40. Cabitza F, Campagner A, Albano D, Aliprandi A, Bruno A, Chianca V, et al. The elephant in the machine: proposing a new metric of data reliability and its application to a medical case to assess classification reliability. Appl Sci. 2020;10:4014.
41. Saluja R, Rai M, Saluja R. Original research article designing new student performance prediction model using ensemble machine learning. J Auton Intell, 6, 2023.
42. Rockman D, Aderibigbe J, Allen-Ile C, Mahembe B, Hamman-Fisher D. Working-class postgraduates' perceptions of studying while working at a selected university. SA J Hum Resource Manage. 2022;20:1115.
43. Santoso A, Retnawati H, Kartianom K, Apino E, Rafi I, Rosyada M. Predicting time to graduation of open university students: an educational data mining study. Open Educ Stud. 2024;6:0216.
44. Seni AJ, Exusper Nemes J. Challenges affecting timely completion of evening postgraduate programs in Tanzania: insights from students' and supervisors' perspectives. Int J Learn Spaces Stud, pp. 1–15, 2025.
45. Injadat M, Moubayed A, Nassif AB, Shami A. Systematic ensemble model selection approach for educational data mining. Knowl Based Syst. 2020;200:105992.
46. Sadqui A, Ertel M, Sadiki H, Amali S. Evaluating machine learning models for predicting graduation timelines in Moroccan universities. Int J Adv Comput Sci Appl, 14, 2023.
47. Molefe D, Tsayang G, Shehu J. Demographic factors influencing graduate programme completion at the university of Botswana, vol. 27, pp. 19–34, 2024.
48. Ruete D, Zavala G, Leal D, Lira PL, Medina LPSM, Aravena ME et al., Early detection of delayed graduation in master's students, in. 2021 ASEE Virtual Annual Conference Content Access, 2021.

Publisher's note