FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 10 No. 23, December, 2026, pp 1 – 9
DOI: https://doi.org/10.33003/fjs-2026-1023-6197

Longitudinal Learning Analytics for Academic Trajectories and Non-Graduation Risk: Evidence from A Nigerian College of Education

***[1]Ismail Adegbola, [1]Sheriffdeen Yinusa [1]Taiwo Olateju-Fajemibola, [2]Oluwatoyin Obansola, [1]Maruf Alimi**

[1]Department of Computer Science, Federal University of Agriculture and Development Studies, Iragbiji, Nigeria
[2]Deptartment of Software Engineering, Federal university of Agriculture and Technology Okeho, Nigeria

*Corresponding authors' email: ismail.adegbola@fuadsi.edu.ng
https://orcid.org/0000-0002-0268-1992

Received: 08 Sept. 2026
Accepted: 06 Nov. 2026
Published: 10 Nov. 2026

ABSTRACT
Student progression and successful programme completion are important concerns in higher education. This study combined longitudinal trajectory analysis and explainable machine learning to examine academic progression and predict recorded non-graduation among Nigeria Certificate in Education (NCE) students. Academic records of 817 students from the 2020–2022 entry cohorts were analysed. Fifteen features derived from three programme years were examined using K-means clustering, while 18 predictors from Years 1 and 2 were used for early prediction. Recorded outcomes comprised 593 graduates and 224 non-graduates. Internal clustering validation favoured two clusters, with non-graduation rates of 87.1% and 12.5% ($\chi^2(1) = 364.68$, $p < .001$; Cramér's V = 0.668). An exploratory four-cluster solution revealed more detailed progression patterns, with non-graduation rates ranging from 5.9% to 92.1%. Logistic Regression, Random Forest, Gradient Boosting and Support Vector Machine were compared using nested five-fold cross-validation. Logistic Regression achieved the highest ROC-AUC (0.9089) and average precision (0.8361), with a Brier score of 0.1085. Bootstrap analysis yielded a 95% confidence interval of 0.8846–0.9327 for ROC-AUC. Coefficient and SHAP analyses identified academic marks, course failures and changes in failure rates as influential predictors. The findings suggest that longitudinal learning analytics can characterise academic trajectories and support earlier identification of students requiring academic assistance.

Keywords: Learning Analytics, NCE Students, Academic Trajectories, K-means Clustering, Non graduation prediction, Logistic Regression

INTRODUCTION

Student progression and successful programme completion remain important concerns for higher institutions. Failure to complete a programme on schedule has consequences for students, institutions, and the efficient use of educational resources (Marzouk, 2025). Consequently, institutions increasingly seek methods for identifying students who may experience learning difficulty early enough to warrant support (Villar & de Andrade, 2024; Vaarma & Li, 2024; Marzouk, 2025). The growth of educational data has created new opportunities for this purpose. Institutional information systems routinely generate records containing course registrations, examination marks, grades, failed courses, course attempts, progression status, and graduation outcomes. When systematically analysed, these records can support learning analytics and predictive modelling without requiring institutions to collect additional information from students.

Learning analytics enables higher-education institutions to use routinely generated educational data to understand student progression, academic performance, and retention. A major advantage is that student progression can be represented as a temporal process. Academic performance may improve, decline, remain stable, or fluctuate, and these patterns may carry information not visible in a single cumulative performance indicator (Lim et al., 2022; Pêgo et al., 2024).

Machine learning has become an increasingly common approach to predicting student dropout, retention, academic failure, and programme completion. Studies by Villar and Robledo Velini de Andrade (2024), Goren et al. (2024), Vaarma and Li (2024), and Osemwegie and Amadin (2023) demonstrate the usefulness of supervised models and institutional academic records for early risk prediction. Systematic reviews also report growing applications of machine learning and deep learning in higher-education dropout prediction (Duro et al., 2026), while Pelima et al. (2024) identified 70 studies examining graduation prediction. However, predictive performance alone does not necessarily explain the academic pathways leading to an outcome. Vaarma and Li (2024) identified accumulated credits and failed courses as important predictors, highlighting the value of academic records for identifying risk.

Academic performance is dynamic: students may maintain stable performance, improve after an initially weak start, experience progressive deterioration, or fluctuate across academic periods. Treating students as static observations may therefore conceal meaningful differences in academic progression. Lim et al. (2022) examined trajectories among more than 4,000 students using dynamic time warping and hierarchical clustering, demonstrating the potential of trajectory patterns to provide information associated with subsequent academic performance. Their study used student-location/semantic trajectories associated with academic performance. Pêgo et al. (2024) examined complex higher-education trajectories involving delayed completion, programme changes, and other deviations from conventional progression pathways. Together, these studies support examining academic progression as a longitudinal process rather than relying exclusively on a single performance indicator.

Trajectory analysis complements binary classification by identifying groups of students with similar academic-performance patterns before comparing those groups with subsequent outcomes. This approach can distinguish persistent difficulty, deterioration, stability, and recovery, thereby providing a more detailed representation of academic progression (Lim et al., 2022; Pêgo et al., 2024). However, trajectory discovery alone does not establish how accurately students at risk can be identified using information available early enough for intervention.

Repeated academic failure may lead to accumulated course requirements and progression difficulties. Accordingly, failed courses, failure rates, course attempts, marks, and changes in these indicators across academic years are relevant to examining academic trajectories and estimating non-graduation risk (Vaarma & Li, 2024). For an early-warning model to support timely intervention, its predictors must also be available at the intended prediction point. Incorporating later academic information can introduce data leakage and inflate apparent predictive performance. This study therefore uses the first three academic years for retrospective trajectory discovery but restricts early prediction to Year-1 and Year-2 information.

Predictive accuracy is not sufficient when machine-learning systems are intended to support educational decisions (Tiukhova et al., 2024). Academic advisers and administrators also need to understand the factors contributing to estimated student risk. Explainable artificial intelligence (XAI) can support this interpretation by providing complementary perspectives on model predictions (Nagy & Molontay, 2024; Mustofa et al., 2025). This study uses standardised Logistic Regression coefficients and SHAP analysis to examine the direction and magnitude of predictor associations and their contributions to individual model predictions. These explanations are intended to support interpretation of the model rather than establish causal relationships.

Existing research demonstrates the value of machine learning for predicting student outcomes and the potential of longitudinal analytics for characterising academic progression. Nevertheless, important methodological gaps remain. First, prediction-focused studies may give limited attention to identifying distinct longitudinal academic pathways and comparing them with completion outcomes. Second, although failed courses and academic performance are established risk indicators in the cited literature, examining changes in annual failure rates, failed-course counts and academic marks across successive programme years may provide additional insight into different progression patterns. Third, trajectory discovery using complete historical records and early prediction using information available at a defined prediction point are not always integrated within a clearly separated analytical framework. Finally, explainability must be combined with appropriate cross-validation procedures and leakage control if predictive findings are to inform institutional decisions responsibly.

To address these gaps, this study integrates three-year longitudinal trajectory discovery, comparison of trajectory groups with recorded graduation outcomes, early prediction of eventual non-graduation using Year-1 and Year-2 information, and model explainability. It examines the academic-performance trajectories identifiable from students' records, how recorded graduation outcomes vary across these trajectories, the predictive performance of four machine-learning classifiers evaluated through nested five-fold cross-validation, and the academic indicators contributing to estimated non-graduation risk. Separating retrospective trajectory discovery from early prediction prevents Year-3 information from entering the early-warning model.

The aim of this study is to develop and evaluate an explainable learning-analytics framework that identifies longitudinal academic trajectories and estimates eventual non-graduation risk among NCE students using institutional academic records.

MATERIALS AND METHODS

Research Design

The study employed a retrospective longitudinal learning-analytics design combining descriptive statistics, unsupervised machine learning, inferential testing, supervised machine learning, and model explainability. Three-year academic records were used to identify and characterise patterns of student progression through trajectory clustering, while Year-1 and Year-2 academic information was used to develop and evaluate models for predicting recorded non-graduation. Four classification algorithms were compared using nested stratified five-fold cross-validation. Logistic Regression was subsequently examined as the primary model for coefficient-based and SHAP-based explainability.

Study Setting and Data Source

Institutional academic records from Oyo State College of Education, Lanlate, Nigeria were used. The source database contained student information and detailed course-level results. The original database contained 1,293 student records. Eligibility produced 817 students from the 2020, 2021, and 2022 entry cohorts with complete Years 1–3 records and established graduation outcomes. Of the 476 excluded records, 431 belonged to entry cohorts outside the study period, while 45 students within the selected cohorts lacked recorded results covering all three programme years. The NCE programme normally spans three academic years, although students with outstanding requirements may take up to five years to complete the programme.

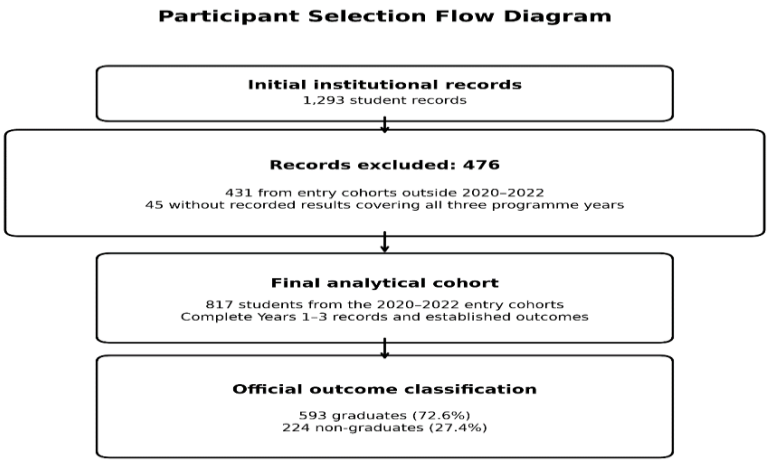


Figure 1: Participant Selection, Exclusion Reasons, And Final Analytical Cohort

Outcome Definition

The supervised outcome was official non-graduation, coded as 1 = non-graduate and 0 = graduate. The final sample contained 593 graduates and 224 non-graduates. Graduation classifications were verified against the institution's Graduating List and Non-Graduating List, which were generated using the institution's established graduation criteria applied to students' academic records. Students who completed their programmes after four or five years were classified as graduates rather than non-graduates. Non-graduation therefore represents the recorded institutional status at the date of ascertainment and does not necessarily imply permanent non-completion.

Data preparation and Features

Student identifiers were normalized before linking records. Course results were ordered by session, semester, and course and aggregated to student-year level. An anomalous mark of 999.99 was treated as invalid. During data preparation, 13 exact duplicate result records were identified and removed. One conflicting result record was identified for the same student-course-session combination and resolved as a data-quality issue after verification against the authoritative institutional record.

Fifteen features were constructed for trajectory analysis. These comprised annual course attempts, failed courses, failure rates and mean marks for Years 1–3, together with three longitudinal measures: the ordinary least-squares slopes of annual mean marks and failure rates across the three programme years, and the sample standard deviation of annual mean marks. The slopes captured the direction of change in academic performance and failure rates, while the standard deviation measured variation in annual mean marks.

Trajectory Identification

Fifteen longitudinal features derived from academic records across Years 1–3 were used to characterise students' academic progression. These features captured annual marks, failure rates, failed-course counts, changes in performance over time, and variability in annual mean marks. Missing feature values were median-imputed, and the features were standardised using z-score scaling. K-means clustering was implemented using the Lloyd algorithm, with 50 initialisations (n_init = 50) and a fixed random seed (random_state = 42). Candidate solutions containing two to six clusters were evaluated using the silhouette coefficient, Calinski–Harabasz index, and Davies–Bouldin index.

Cluster stability was assessed for the two- and four-cluster solutions using 100 subsamples, each comprising 80% of students drawn without replacement. The adjusted Rand index (ARI) was calculated by comparing assignments for shared students with the corresponding full-data solution. Imputation and standardisation were performed before subsampling; therefore, the stability assessment did not capture variability arising from refitting these preprocessing steps.

Trajectory Comparison

Trajectory groups were compared with official graduation outcomes using Pearson's chi-square test of independence, with Cramér's V used to assess the strength of association. The analysis was conducted separately for the two-cluster primary solution and the four-cluster exploratory solution.

Early Prediction Dataset and Predictors

The early-prediction dataset contained 18 numeric predictors derived from Year-1 and Year-2 academic records, including course attempts, failed courses, failure rates, marks, numbers of passes and unique courses, and changes in these indicators between the two years. Student identifiers, cohort identifiers, official-status fields, and the target outcome were excluded from the predictor set. No Year-3 variable or three-year trajectory-cluster assignment was used as a predictor, ensuring that the prediction model relied on information available from the first two academic years.

Prediction Models

Four supervised classifiers were evaluated: Logistic Regression, Random Forest, Gradient Boosting and Support Vector Machine (SVM). Each was implemented using a scikit-learn pipeline, with median imputation applied to missing predictor values and standardisation included for Logistic Regression and SVM. Tree-based models were evaluated without feature standardisation.

Hyperparameter optimisation was conducted using GridSearchCV within the five-fold stratified inner cross-validation. The search spaces comprised 20 Logistic Regression configurations (C = 0.01–100, two class-weight options and two solvers), 16 Random Forest configurations (300 or 500 trees, minimum leaf size of 1 or 2, two maximum-feature settings and two class-weight options), 16 Gradient Boosting configurations (100 or 200 estimators, learning rates of 0.05 or 0.1, maximum depths of 2 or 3 and minimum leaf sizes of 1 or 2), and six SVM configurations (C = 0.1, 1 or 10, RBF kernel and two class-weight options). ROC-AUC was

used to select the best configuration for each classifier. The selected estimator was refitted on the complete outer-training fold and evaluated on the corresponding held-out fold.

Validation and Performance Evaluation

Predictive performance was evaluated using nested stratified five-fold cross-validation with shuffling and a fixed random seed (42). The outer loop provided held-out performance estimates, while the inner loop supported hyperparameter optimisation and classification-threshold selection within each outer training partition. Out-of-fold probabilities from the five outer folds were pooled to estimate overall performance.

Classification thresholds were selected using inner out-of-fold predictions by maximising the F1-score and then applied to the corresponding outer test predictions. Model performance was assessed using ROC-AUC, average precision, accuracy, precision, sensitivity, specificity, F1-score and Brier score. Held-out test observations were not used for model or threshold selection.

Explainability

Model explainability was assessed using standardised Logistic Regression coefficients and SHAP LinearExplainer. Coefficients were summarised across the five outer folds, while SHAP values were calculated for held-out observations using training-fold preprocessing. Mean absolute SHAP values were used to assess global feature importance. Both methods were interpreted as explanations of model behaviour rather than evidence of causal relationships.

Statistical Analysis

Descriptive statistics summarised student characteristics, academic performance and graduation outcomes. Associations between trajectory clusters and graduation status were examined using Pearson's chi-square test and Cramér's V, with 95% Wilson confidence intervals for cluster-specific non-graduation rates. Predictive performance uncertainty for the primary Logistic Regression model was assessed using 2,000 stratified bootstrap resamples of students' pooled out-of-fold predictions. Percentile-based 95% confidence intervals were calculated for ROC-AUC, average precision and Brier score. As models were not retrained during bootstrapping, these intervals reflect uncertainty conditional on the saved predictions rather than the full model-development process. Statistical significance was set at $p < .05$ (two-sided).

RESULTS AND DISCUSSION

Student Cohort Characteristics and Outcome Distribution

The study dataset comprised 817 students from the 2020, 2021, and 2022 entry cohorts, selected from an institutional database containing 1,293 student records. The final analytical sample consisted of 593 graduates (72.6%) and 224 non-graduates (27.4%). The 2020 cohort contributed 342 students, comprising 240 graduates and 102 non-graduates. The 2021 cohort comprised 275 students, including 194 graduates and 81 non-graduates, while the 2022 cohort contained 200 students, of whom 159 were graduates and 41 were non-graduates. The distribution indicates that non-graduation was present across all three entry cohorts, although the proportion varied. Non-graduation rates were 29.8%, 29.5% and 20.5% for the 2020, 2021 and 2022 cohorts, respectively

Longitudinal Academic Trajectory Clustering

K-means clustering was applied to 15 longitudinal features, evaluating solutions of two to six clusters using inertia, silhouette coefficient, Calinski–Harabasz and Davies–Bouldin indices. The two-cluster solution showed the best internal validation, with a silhouette coefficient of 0.428, Calinski–Harabasz index of 384.21, Davies–Bouldin index of 1.213 and inertia of 8,328.69.

Stability assessment supported this solution, with a mean ARI of 0.969 (SD = 0.029; median = 0.979; range = 0.850–1.000), compared with 0.905 (SD = 0.092; median = 0.934; range = 0.596–0.996) for the four-cluster solution. Accordingly, the two-cluster solution was retained for the primary analysis, while the four-cluster solution was examined exploratorily to provide more detailed academic progression profiles. Clusters were interpreted as descriptive groups rather than causal categories.

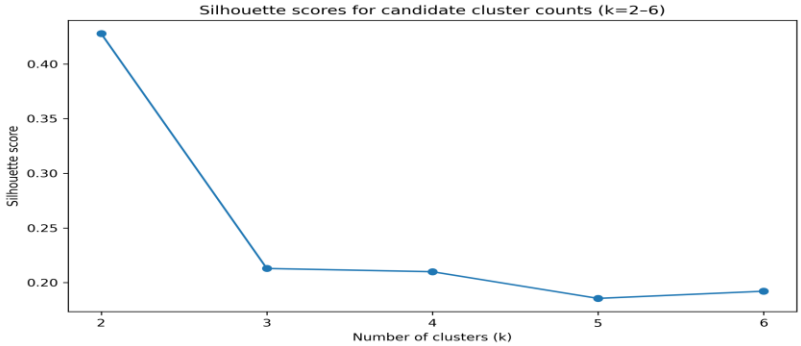


Figure 2: Silhouette Scores for Candidate Cluster Counts (k=2–6); the Highest Score Occurred at k=2

Cluster Profiles and Academic Progression Patterns

The exploratory four-cluster solution provided a more detailed representation of academic performance and failure patterns across the three years. The clusters differed in their mean marks, course failure rates, numbers of failed courses, and recorded non-graduation proportions.

Cluster 0 contained 338 students. Their mean marks were 45.91, 46.30, and 47.41 across Years 1, 2, and 3, respectively. Their failure rates were 7.26%, 6.80%, and 1.68%, while the mean numbers of failed courses were 2.03, 2.22, and 0.31. This pattern indicates generally moderate academic performance, with a substantial reduction in course failures by the third year.

Cluster 1 comprised 63 students and showed a declining mean-mark pattern, from 43.61 in Year 1 to 40.26 in Year 2 and 38.45 in Year 3. Its failure rates increased from 11.59%

in Year 1 to 21.18% in Year 2 and 24.85% in Year 3. The mean numbers of failed courses were 2.83, 6.25, and 3.70, respectively. This cluster represented a group with persistently weak and deteriorating academic performance.

Cluster 2 contained 93 students. The mean marks were 36.91, 36.50, and 40.88 across the three years. Failure rates were 33.03%, 33.09%, and 15.52%, while the mean numbers of failed courses were 8.73, 10.95, and 2.84. The pattern indicates substantial academic difficulty during the first two years, followed by an improvement in the third year, although the earlier failure burden remained considerable.

Cluster 3 comprised 323 students and exhibited the strongest academic performance. Its mean marks were 56.59, 56.91, and 56.45 across the three years. Failure rates remained very low at 0.53%, 0.65%, and approximately 0.04%, while the mean numbers of failed courses were 0.15, 0.20, and approximately 0.01. This cluster therefore represented students with consistently strong academic performance and very few recorded course failures.

The four-cluster profiles demonstrate that students followed different academic progression patterns, ranging from consistently strong performance to persistent or changing patterns of academic difficulty.

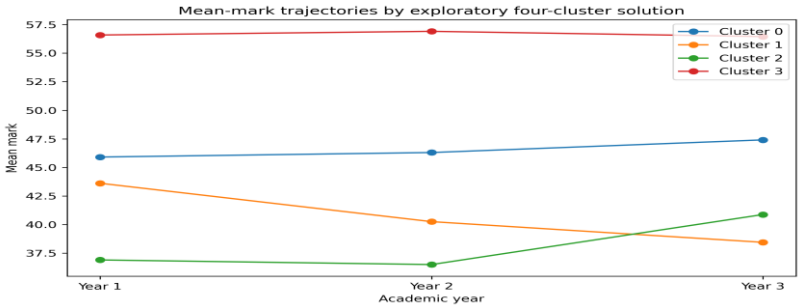


Figure 3: Mean Academic Marks Across Years 1–3 for the four Exploratory Clusters

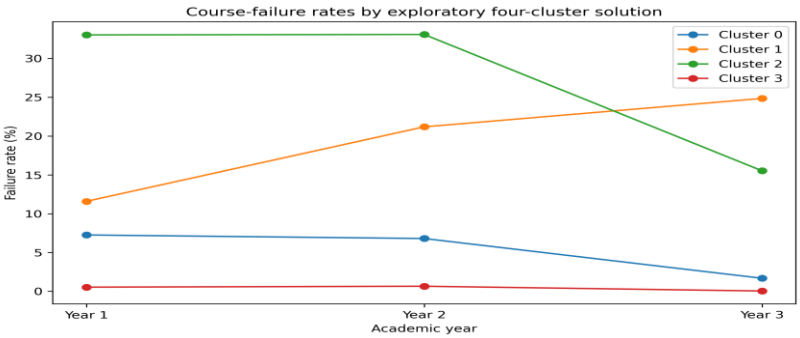


Figure 4: Course Failure Rates Across Years 1–3 for the four Exploratory Clusters

Association Between Trajectory Clusters and Graduation Status

The primary two-cluster solution showed a significant association with recorded graduation status, $\chi^2(1, N = 817) = 364.68$, $p < .001$, Cramér's V = 0.668. Cluster 0 contained 163 students (21 graduates and 142 non-graduates), with a non-graduation rate of 87.1% (95% Wilson CI: 81.1%–91.4%). Cluster 1 contained 654 students (572 graduates and 82 non-graduates), with a rate of 12.5% (95% Wilson CI: 10.2%–15.3%).

The exploratory four-cluster solution also showed a significant association, $\chi^2(3, N = 817) = 377.87$, $p < .001$, Cramér's V = 0.680. Non-graduation rates were 19.8% (67/338), 92.1% (58/63), 86.0% (80/93) and 5.9% (19/323) for Clusters 0–3, respectively. The two-cluster solution remained the primary result based on its stronger internal validation and stability, while the four-cluster solution provided additional descriptive detail. These associations do not establish a causal relationship between cluster membership and non-graduation.

Table 1: Cluster Composition and Recorded Non-Graduation Rates for The Primary Two-Cluster and Exploratory Four-Cluster Solutions

Solution	Cluster	n	Graduates	Non-graduates	Non-graduation rate	95% Wilson CI
Primary k=2	0	163	21	142	87.1%	81.1%–91.4%
Primary k=2	1	654	572	82	12.5%	10.2%–15.3%
Exploratory k=4	0	338	271	67	19.8%	15.9%–24.4%
Exploratory k=4	1	63	5	58	92.1%	82.7%–96.6%
Exploratory k=4	2	93	13	80	86.0%	77.5%–91.6%
Exploratory k=4	3	323	304	19	5.9%	3.8%–9.0%

Table 2: Overall Cluster Outcome Association Tests

Solution	N	χ^2	df	p-value	Cramér's V
k=2 (primary)	817	364.68	1	<.001	.668
k=4 (exploratory)	817	377.87	3	<.001	.680

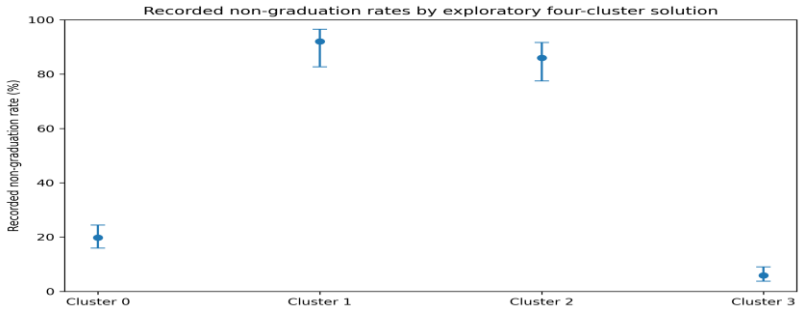


Figure 5: Recorded Non-graduation Rates with 95% Wilson Confidence Intervals by Exploratory Four-cluster Solution

Nested Cross-Validation Predictive Model Evaluation

The four classifiers were evaluated using 817 students and 18 academic predictors derived from Years 1 and 2. Based on pooled outer-fold predictions, Logistic Regression achieved the highest ROC-AUC (0.909) and average precision (0.836), followed by Random Forest (0.906; 0.826), Gradient Boosting (0.905; 0.827) and Support Vector Machine (0.899; 0.778), respectively.

Brier scores were 0.1085 for Logistic Regression, 0.0987 for Random Forest, 0.1004 for Gradient Boosting and 0.1037 for Support Vector Machine. Random Forest recorded the lowest Brier score.

Model Comparison and Classification Performance

Classification thresholds were selected using inner training-fold predictions and applied to the corresponding outer test predictions. The pooled classification results are presented in Table 3. Support Vector Machine achieved the highest accuracy (0.867), precision (0.783), specificity (0.926) and F1-score (0.745). However, Logistic Regression achieved the highest ROC-AUC and average precision. Logistic Regression was retained for explainability analysis because of its stronger discrimination performance, while the other classifiers served as comparative models. This selection does not imply superior performance across all evaluation metrics.

Table 3: Nested Cross-Validation Performance of the Four Predictive Models

Model	ROC-AUC	Average precision	Brier score	Accuracy	Precision	Sensitivity	Specificity	F1-score
Logistic Regression	0.909	0.836	0.1085	0.847	0.722	0.719	0.895	0.720
Random Forest	0.906	0.826	0.0987	0.851	0.730	0.723	0.899	0.726
Gradient Boosting	0.905	0.827	0.1004	0.857	0.758	0.701	0.916	0.729
Support Vector Machine	0.899	0.778	0.1037	0.867	0.783	0.710	0.926	0.745

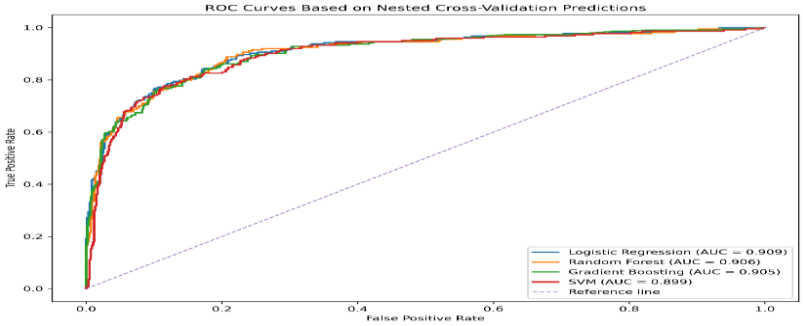


Figure 6: ROC Curves Based On Pooled Outer-Fold Predictions for The Four Classification Models

Bootstrap Uncertainty Analysis

Uncertainty estimates for the primary Logistic Regression model are presented in Table 4. The model achieved an ROC-AUC of 0.9089, average precision of 0.8361 and Brier score of 0.1085, with their corresponding 95% bootstrap confidence intervals.

Table 4: Bootstrap Uncertainty Estimates for Logistic Regression

Metric	Estimate	Lower 95% CI	Upper 95% CI
ROC-AUC	0.9089	0.8846	0.9327
Average Precision	0.8361	0.7993	0.8733
Brier Score	0.1085	0.0968	0.1202

Logistic Regression Coefficient Analysis

Standardised Logistic Regression coefficients showed that Year 1 median mark (−0.4916, SD = 0.5475), Year 1 failed courses (0.4175, SD = 0.2594), Year 2 failure rate (0.3962, SD = 0.2122), Year 1 passes (−0.3473, SD = 0.1856), change in failure rate (0.3147, SD = 0.3238) and change in mean mark (−0.2694, SD = 0.2414) had consistent coefficient directions across all five folds. Year 2 failed courses and Year 1 failure rate also showed consistent positive directions. These findings indicate that previous academic performance, course failures and changes in performance contributed to the model's predictions, although coefficient magnitudes varied across folds. The coefficients represent predictive associations, not causal effects.

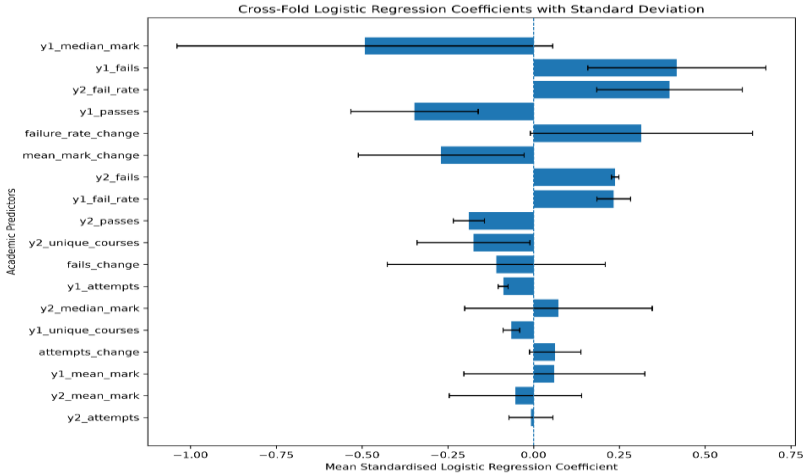


Figure 7: Standardised Logistic Regression Coefficients across the Five Outer Folds, Showing mean Coefficients and Variability

SHAP-Based Model Explainability

SHAP analysis identified Year 1 median mark (0.4040), Year 1 failed courses (0.3082), Year 2 failure rate (0.2798), Year 1 passes (0.2491), change in mean mark (0.2044) and change in failure rate (0.1932) as the six leading contributors to model predictions. The SHAP summary plot (Figure 8) showed that higher Year 1 median marks and more passes generally contributed towards lower predicted non-graduation risk, whereas greater course failures and higher failure rates contributed towards increased risk. These findings were broadly consistent with the standardised Logistic Regression coefficients. SHAP contributions represent model output on the log-odds scale and do not establish causal effects.

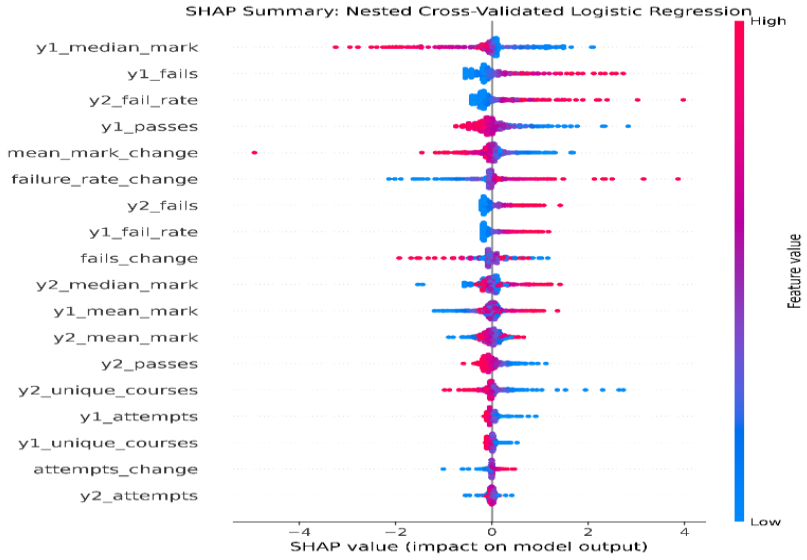


Figure 8: SHAP summary Plot of the Nested Cross-Validated Logistic Regression Model

Discussion

This study integrated longitudinal trajectory analysis and explainable machine learning to examine academic progression and predict recorded non-graduation among NCE students. The findings address four areas: academic progression patterns, their association with graduation outcomes, predictive performance and the academic indicators contributing to model predictions.

First, the trajectory analysis identified distinct patterns of academic progression. The two-cluster solution provided the strongest internal validation and stability, while the exploratory four-cluster solution revealed more differentiated academic profiles, including students with sustained strong performance and others experiencing persistent or changing academic difficulties. These findings suggest that academic progression cannot always be adequately characterised by isolated examination results. Examining changes in academic performance and course failures across successive years provides additional insight into students' progression patterns and potential support needs.

Second, substantial associations were observed between trajectory clusters and recorded graduation status. The non-graduation rates varied considerably across the primary two-cluster and exploratory four-cluster solutions. This indicates that students with different longitudinal academic profiles also experienced markedly different graduation outcomes. However, these data-derived associations do not establish causation or demonstrate independent prospective prediction.

Third, the predictive modelling results support previous research on machine learning for student dropout and academic success prediction (Villar & Robledo Velini de Andrade, 2024; Vaarma & Li, 2024). Logistic Regression achieved the highest ROC-AUC (0.9089) and average precision (0.8361), although Random Forest recorded the lowest Brier score, while SVM achieved the highest accuracy and F1-score. Logistic Regression was therefore selected for explainability based on its discrimination and precision-recall performance. These findings suggest that routinely collected academic records can support the identification of students at risk of non-graduation by the end of their second academic year. Nevertheless, external validation is necessary before operational deployment.

Finally, standardised Logistic Regression coefficients and SHAP analysis identified Year 1 median mark, Year 1 failed courses, Year 2 failure rate and Year 1 passes among the leading predictors. Changes in mean marks and failure rates also contributed to model predictions. These findings are broadly consistent with Vaarma and Li (2024), who identified failed courses and accumulated credits as important indicators of student outcomes. The agreement between the two explanation methods provides complementary evidence about the academic indicators represented in the model. However, these explanations describe predictive associations rather than causal relationships, and their generalisability to other institutional settings remains to be established.

The findings demonstrate the complementary value of longitudinal trajectory analysis and explainable machine learning for examining academic progression and estimating non-graduation risk. The approach provides a basis for evidence-informed student-support strategies, subject to external validation and responsible interpretation within different institutional contexts.

Limitations and Future Research

The findings are limited by the single-institution setting and absence of external validation. Non-academic factors influencing student progression were not included. Coefficient and SHAP analyses indicate predictive associations rather than causal relationships.

CONCLUSION

This study combined longitudinal learning analytics and explainable machine learning to examine academic progression and predict recorded non-graduation among National Certificate in Education (NCE) students. The trajectory analysis identified distinct academic progression patterns, while predictive modelling using Year 1 and Year 2 records showed that Logistic Regression achieved the highest ROC-AUC (0.909) and average precision (0.836) among the four evaluated models. Coefficient and SHAP analyses identified academic performance, course failures and changes in failure rates as important contributors to model predictions. The findings demonstrate the potential of combining trajectory analysis and explainable predictive modelling to support early identification of students requiring academic assistance. Future research should extend the approach to multiple institutions, incorporate relevant non-academic factors, and conduct prospective evaluations to determine whether prediction-informed interventions improve student progression and programme completion.

ACKNOWLEDGEMENTS

The author gratefully acknowledges Oyo State College of Education, Lanlate, Oyo State, Nigeria, for providing access to the institutional academic records used in this study.

ETHICS

This study used institutional academic records obtained with a permission waiver. The dataset used for analysis was handled as institutional research data, and personally identifying information was not included in the reported analyses.

DATA AVAILABILITY

The dataset contains educational records and is therefore not publicly available in identifiable form.

REFERENCES

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), Article 164. https://doi.org/10.3390/computers15030164

Goren, O., Cohen, L., & Rubinstein, A. (2024). Early prediction of student dropout in higher education using machine learning models. In Proceedings of the 17th International Conference on Educational Data Mining. https://educationaldatamining.org/EDM2024/proceedings/2024.EDM-short-papers.32/

Lim, H., Kim, S., Chung, K.-M., Lee, K., Kim, T., & Heo, J. (2022). Is college students' trajectory associated with academic performance? Computers & Education, 178, Article 104397. https://doi.org/10.1016/j.compedu.2021.104397

Marzouk, O. A. (2025). Benchmarking retention, progression, and graduation rates in undergraduate higher education across different time windows. Cogent Education, 12(1), Article 2498170. https://doi.org/10.1080/2331186X.2025.2498170

Mustofa, S., Emon, Y. R., Mamun, S. B., Akhy, S. A., & Ahad, M. T. (2025). A novel AI-driven model for student

dropout risk analysis with explainable AI insights. Computers and Education: Artificial Intelligence, 8, Article 100352. https://doi.org/10.1016/j.caeai.2024.100352

Nagy, M., & Molontay, R. (2024). Interpretable dropout prediction: Towards XAI-based personalized intervention. International Journal of Artificial Intelligence in Education, 34, 274–300. https://doi.org/10.1007/s40593-023-00331-8

Osemwegie, E. E., & Amadin, F. I. (2023). Student dropout prediction using machine learning. FUDMA Journal of Sciences, 7(6), 347–353. https://doi.org/10.33003/fjs-2023-0706-2103

Pêgo, J. P., Miguéis, V. L., & Soeiro, A. (2024). Students' complex trajectories: Exploring degree change and time to degree. International Journal of Educational Technology in Higher Education, 21, Article 8. https://doi.org/10.1186/s41239-024-00438-5

Pelima, L. R., Sukmana, Y., & Rosmansyah, Y. (2024). Predicting university student graduation using academic performance and machine learning: A systematic literature review. IEEE Access, 12, 23451–23465. https://doi.org/10.1109/ACCESS.2024.3361479

Tiukhova, E., Vemuri, P., López Flores, N., Islind, A. S., Óskarsdóttir, M., Poelmans, S., Baesens, B., & Snoeck, M. (2024). Explainable learning analytics: Assessing the stability of student success prediction models by means of explainable AI. Decision Support Systems, 182, Article 114229. https://doi.org/10.1016/j.dss.2024.114229

Vaarma, M., & Li, H. (2024). Predicting student dropouts with machine learning: An empirical study in Finnish higher education. Technology in Society, 76, Article 102474. https://doi.org/10.1016/j.techsoc.2024.102474

Villar, A., & de Andrade, C. R. V. (2024). Supervised machine learning algorithms for predicting student dropout and academic success: A comparative study. Discover Artificial Intelligence, 4, Article 2. https://doi.org/10.1007/s44163-023-00079-z

©2026 This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license viewed via https://creativecommons.org/licenses/by/4.0/ which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is cited appropriately.