Cureus Journal of Computer Science
https://doi.org/10.7759/s44389-026-00347-w

Original Article

Leakage-Safe Ensemble Learning for Early Prediction of Student Degree Outcomes in a Nigerian University

Olumayokun A. Tanimonure [1], Adekunle O. Afolabi [2, 3], Damilola E. Ogunniyi [4]

1. *Department of Information Systems, Obafemi Awolowo University, Ile-Ife, NGA*
2. *Department of Computer Science and Engineering, University at Buffalo, Buffalo, USA*
3. *Department of Computer Science and Engineering, Obafemi Awolowo University, Ile-Ife, NGA*
4. *Information Technology, Osun State University, Osogbo, NGA*

Received: May 11, 2026 | Review began: June 22, 2026 | Review ended: September 17, 2026 | Published: October 30, 2026

© **Copyright** 2026
This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

This paper proposes a leakage-controlled ensemble learning framework for the staged prediction of undergraduates' final degree class from longitudinal student records of a Nigerian university. The application context is early warning/advising systems where there is a need to identify students suitable for interventions prior to graduation. Anonymous academic history data of students who graduated from a five-year Computer Science and Engineering program were prepared into three-, four- and five-year prediction points. Potential predictor attributes comprised students' year-wise cumulative grade point average (CGPA), average scores, cumulative registered units and academic standing. Leakage was controlled by restricting information available for early predictions to exclude students' final CGPA, final degree class, and attributes from future academic years. Random Forest (RF) and Extreme Gradient Boosting (XGBoost) classifiers were evaluated against each other under a validation framework designed for both information leakage and class imbalance. Metrics for evaluation include accuracy, macro- and weighted-F1, balanced accuracy and confusion matrix error distributions instead of accuracy alone. Preliminary findings indicate that overall, XGBoost yielded better and more consistent classification performances than RF especially at the beginning and end of the academic program, while RF had comparable performances in the intermediate years. Ensemble learning methods could be utilized for data-informed student advising if prediction outcomes are used as intervention signals instead of categorical academic decisions. We share insights from our staged prediction framework, leakage control heuristics and an in-house deployment path for educational data mining in developing world universities.

Categories: Data Engineering, Machine Learning (ML), Data Visualization

Keywords: ensemble learning, student outcomes, degree classification, learning analytics, educational data mining, random forest, xgboost, class imbalance, academic advising, modeling framework

Introduction

Forecasting students' academic achievement is a central priority in higher education because institutions require actionable, timely empirical evidence to support academic advising, student retention strategies, and institutional resource allocation. Learning analytics and educational data mining (EDM) can process historical student records to identify at-risk undergraduates prior to the compounding of academic failure [1-4]. In Nigerian universities, degree classification strongly influences graduate employability, competitive scholarship opportunities, and postgraduate

admissions. This institutional reality requires predictive models that can forecast final degree classifications early in students' academic careers, enabling targeted, human-led academic interventions before graduation outcomes become immutable.

Under the National Universities Commission regulatory framework in Nigeria, undergraduate academic performance is evaluated on a five-point cumulative grade point average (CGPA) scale derived from course credit unit weightings and examination scores. Final degree classifications are awarded at graduation across five standardized tiers: First Class Honors (CGPA 4.50-5.00), Second Class Honors Upper Division (3.50-4.49), Second Class Honors Lower Division (2.40-3.49), Third Class Honors (1.50-2.39), and Pass or Fail (below 1.50). When institutional predictions are delayed until late in the degree program, they forfeit their utility as early warning signals. Consequently, an urgent research imperative is to determine how early an accurate and statistically reliable degree class forecast can be established, and to quantify how predictive fidelity evolves as progressive longitudinal academic records accumulate.

While earlier studies have employed conventional statistical techniques, linear models, and standalone machine learning (ML) algorithms to predict student attrition, course completions, and grade point trajectories [5-11], ensemble learning methodologies offer distinct theoretical and empirical advantages for structured, tabular educational datasets. Tabular student progression records typically exhibit complex nonlinear relationships, multicollinearity among semester GPAs, and moderate sample cohort sizes where deep neural networks frequently overfit and lack appropriate inductive biases [12,13]. Conversely, single baseline estimators - such as individual decision trees - often suffer from high variance and structural instability when trained on real-world academic cohorts. Ensemble architectures address these limitations by aggregating collections of base estimators: Random Forest (RF) leverages bootstrap aggregation (bagging) with random feature subspacing to decorrelate individual trees and reduce overall variance [14], whereas Extreme Gradient Boosting (XGBoost) sequentially builds shallow trees by optimizing custom objective functions via second-order Taylor expansions and explicit regularization to minimize both bias and variance [15]. In this study, these two canonical ensemble paradigms are benchmarked against nonensemble baselines to establish their comparative effectiveness on longitudinal academic data.

A central methodological pillar of this investigation is the formulation of a leakage-safe, fairness-staged prediction framework. Academic data mining often suffers from subtle forms of target leakage, where variables that inherently encode the final degree class (such as final-year cumulative GPA or graduation standing) or features from future semesters are inadvertently included during model training, leading to unrealistically optimistic performance estimates. The experimental design enforces strict temporal inference boundaries that systematically eliminate all future-year features and target-defining fields from early prediction checkpoints. Furthermore, acknowledging that undergraduate degree distributions are heavily skewed toward middle classifications, the validation protocol applies synthetic oversampling (Synthetic Minority Over-sampling Technique (SMOTE)) strictly within training folds, evaluates models using imbalance-resilient metrics (macro-F1 and balanced accuracy), and explains model predictions using post-hoc feature importance and SHapley Additive exPlanations (SHAP) analyses.

Research aim and questions

This study aims to design, evaluate, and interpret leakage-safe ensemble ML models to predict undergraduate degree outcomes from longitudinal academic data. The study addresses five core research questions. First, to what extent can final degree classifications be accurately forecast across progressive academic milestones from Year 1 to Year 5? Second, do tree-based ensemble architectures (RF and XGBoost) demonstrate statistically significant predictive superiority over standard nonensemble baselines (Multinomial Logistic Regression and Single Decision Trees) when evaluated on imbalanced academic data? Third, what specific academic features and progression dynamics exert the strongest influence on model predictions at each academic milestone? Fourth, at which academic stage do models transition from early warning indicators to high-certainty outcome confirmation? Fifth, how can data leakage controls, cross-validation protocols, and model interpretability be operationalized into a responsible, human-supervised academic advising system in higher education?

How to cite this article:
Tanimonure O A, Afolabi A O, Ogunniyi D E (October 30, 2026) Leakage-Safe Ensemble Learning for Early Prediction of Student Degree Outcomes in a Nigerian University. Cureus J Comput Sci 3 : es44389-026-00347-w. DOI https://doi.org/10.7759/s44389-026-00347-w

Contribution of the study

The primary contribution of this research is a validated, leakage-controlled staged prediction methodology for final degree classifications grounded in longitudinal academic records from an African higher education institution (Obafemi Awolowo University (OAU), Nigeria). Methodologically, the study provides formal temporal feature-exclusion rules, documents a leak-free cross-validation oversampling pipeline, empirically demonstrates the need for ensemble methods over simpler baselines, and introduces SHAP-based feature-importance interpretability across academic progression checkpoints. Practically, the study articulates an actionable advising deployment architecture that translates multiclass predictive probabilities into proactive, ethical academic support interventions.

Materials And Methods

Study context and data source

The empirical investigation utilized retrospective, anonymized longitudinal undergraduate academic records obtained from the Department of Computer Science and Engineering at OAU, Ile-Ife, Nigeria. The dataset comprises a total of 277 unique graduating students enrolled in a five-year Bachelor of Science (B.Sc.) degree program across six graduating cohorts spanning the 2015/2016 through 2020/2021 academic sessions. At the raw transactional level, the data encompasses 4,083 semester course-level enrollment records containing course codes, credit unit values, continuous assessment (CA) scores, end-of-semester examination marks, composite scores, and semester grade points. These records were systematically aggregated into student-level annual milestones reflecting longitudinal academic progression. The primary prediction target is the officially conferred final degree classification awarded by the University Senate upon graduation. All personal identifiers, including matriculation numbers and names, were masked prior to data processing.

Ethical considerations and data governance

Ethical clearance and institutional permission to utilize anonymized historical academic records for educational research were secured. Direct student identifiers were scrubbed and replaced with non-traceable synthetic indices before feature extraction. Only academic attributes essential for evaluating progression were retained. In alignment with ethical artificial intelligence (AI) principles and the Nigerian Data Protection Regulation, the predictive system is strictly designed as an advisory decision support mechanism for departmental academic advisors rather than an automated grading, screening, or disciplinary instrument [16].

Prediction outputs are intended to support academic advising rather than replace advisor judgement. Advisors should interpret predictions alongside academic context, student explanations and available support options because predictive accuracy alone does not guarantee explanatory usefulness or responsible interpretation [16].

Dataset organization and prediction points

To emulate real-world institutional deployment, the modeling pipeline was structured as a sequential staged prediction problem. Five distinct milestone datasets were created, where each stage incorporates only the cumulative academic history accessible at the conclusion of that specific academic year (Year 1 through Year 5). This temporal partitioning enables direct empirical comparison between early warning capability and late-stage outcome confirmation. Table *1* summarizes the characteristics of the staged academic datasets.

Dataset	Academic years included	Records	Features	Decision support role
5-Year_Year1_dataset	Year one only	277	7	Early warning
5-Year_Year2_dataset	Years one to two	277	11	Early to intermediate warning

5-Year_Year3_dataset	Years one to three	275	15	Mid-program monitoring
5-Year_Year4_dataset	Years one to four	219	19	Late-stage monitoring
5-Year_Year5_dataset	Years one to five	241	23	Graduation-stage confirmation

TABLE 1: Summary of staged academic datasets used in the study

Feature engineering and leakage control

Feature engineering focused on capturing multidimensional academic trajectories. Candidate predictor variables were grouped into four families: (i) annual and cumulative grade point averages (CGPA_Y1 through CGPA_Y5), (ii) cumulative credit units registered and passed, (iii) mean course examination scores per academic level, and (iv) longitudinal progression velocity, operationalized as the delta-CGPA between consecutive academic years (e.g., delta-CGPA_Y2 = CGPA_Y2 - CGPA_Y1). The target variable was integer-encoded for multiclass classification across the awarded degree classes: First Class (0), Second Class Upper (1), Second Class Lower (2), Third Class (3), and Pass/Fail (4). The encoded target was strictly withheld from predictor matrices.

Data leakage prevention constitutes the foundational architectural constraint of the experimental design. In educational data mining, leakage manifests when models inadvertently exploit features that either directly define the outcome or contain information chronologically unavailable at the target prediction point. Final graduation CGPA and final degree standing were strictly prohibited from all early-stage predictor sets (Years 1 to 4) because they are deterministic proxies for the target variable. Furthermore, all downstream academic records (e.g., Year 3 course outcomes) were rigorously masked when training models for preceding milestones (e.g., Year 1 and Year 2). Table *2* delineates the temporal feature inclusion rules enforced across all experimental stages.

Prediction point	Permitted predictor family	Variables excluded to avoid leakage
Year one	Year one cumulative grade point average, total units and average score	Degree class, encoded class, final cumulative grade point average and year two to year five variables
Year two	Year one to year two cumulative grade point average, total units, average score and cumulative grade point average change	Degree class, encoded class, final cumulative grade point average and year three to year five variables
Year three	Year one to year three academic and progression variables	Degree class, encoded class, final cumulative grade point average and year four to year five variables
Year four	Year one to year four academic and progression variables	Degree class, encoded class, final cumulative grade point average and year five variables

How to cite this article:
Tanimonure O A, Afolabi A O, Ogunniyi D E (October 30, 2026) Leakage-Safe Ensemble Learning for Early Prediction of Student Degree Outcomes in a Nigerian University. Cureus J Comput Sci 3 : es44389-026-00347-w. DOI https://doi.org/10.7759/s44389-026-00347-w

Year five	Year one to year five academic variables, excluding direct label fields	Degree class and encoded class as predictors

TABLE 2: Leakage-safe feature use rule by prediction point

Data preprocessing

Data preprocessing encompassed transcript validation, duplicate detection, categorical encoding, and continuous feature scaling. Records were scrutinized for consistency across academic sessions to ensure accurate credit unit totals and grade point arithmetic. Missing continuous assessment marks - occurring in historical records where only composite examination marks were logged - were imputed using median continuous feature values computed exclusively within training folds. Continuous variables (CGPA, average scores, registered units) were scaled using standard z-score normalization, with scaling parameters fitted exclusively on training subsets to prevent distribution leakage.

Class imbalance handling

Undergraduate degree award distributions in higher education are inherently imbalanced. In the historical cohort analyzed in this study, the majority of graduates attained Second Class Lower (approx. 51%) and Second Class Upper (approx. 39%), whereas First Class (approx. 4%) and Third Class or Pass/Fail (approx. 6%) constituted severe minority classes. Standard ML algorithms trained on such skewed distributions tend to exhibit pronounced majority-class bias, maximizing raw accuracy while failing to identify academically vulnerable minority-class students who require timely intervention. To counter this distribution skew, the SMOTE [17] was applied to generate synthetic feature vectors along line segments connecting k-nearest minority neighbors. Crucially, SMOTE was embedded strictly within the training fold of each cross-validation iteration following the train-validation split. Applying SMOTE to the pooled dataset prior to cross-validation would allow synthetic information from validation cases to contaminate training sets, resulting in data leakage and artificially inflated performance estimates.

Model development, benchmark baselines, and hyperparameter optimization

To establish a rigorous empirical foundation and evaluate the specific advantages of ensemble learning on longitudinal academic data, four predictive architectures comprising two nonensemble baselines and two tree-based ensemble methods were implemented and compared:

1. Multinomial Logistic Regression (Baseline): A standard linear multiclass classifier with L2 regularization (C = 1.0) and cross-entropy loss, establishing the benchmark for linear separability among degree classes.

2. Single Decision Tree (Baseline): A standalone CART decision tree using Gini impurity splitting and a maximum depth of five (max_depth = 5), serving as the nonaggregated tree baseline to isolate the exact variance reduction and boosting gains achieved by ensemble architectures.

3. Random Forest (Ensemble): A bagging-based ensemble comprising 100 decorrelated decision trees, utilizing bootstrap resampling and random feature subspacing (square root of total features) with balanced class weighting to reduce model variance [10].

4. Extreme Gradient Boosting/XGBoost (Ensemble): A scalable gradient-boosted decision tree ensemble that iteratively fits shallow trees to pseudo-residuals using second-order Taylor expansions of the multi-class log-loss objective, incorporating L1/L2 leaf weight regularization [11].

Hyperparameter Optimization Protocol: Hyperparameters were systematically optimized using a repeated 5-fold Stratified Cross-Validation grid search over candidate parameter grids: tree depth in {3, 5, 7, 10}, learning rate eta in {0.01, 0.05, 0.1, 0.2}, number of estimators in {50, 100, 200}, and min_samples_split in {2, 5, 10}. A constrained maximum tree depth of max_depth = 5 and learning rate eta = 0.1 emerged as optimal for XGBoost, while max_depth = 5 with 100

estimators and balanced class weighting was optimal for RF. Constrained depth serves as an essential structural regularizer that prevents individual trees from memorizing idiosyncratic noise in moderate-sized tabular cohorts (N = 277). Table *3* summarizes the finalized model configurations and their theoretical rationales.

Model architecture	Hyperparameter and architecture settings	Theoretical and empirical rationale
Multinomial Logistic Regression (Baseline)	Multiclass multinomial formulation, L2 regularization (C = 1.0), lbfgs solver, max_iter = 1,000	Linear benchmark baseline to quantify linear separability of degree classes
Single Decision Tree (Baseline)	CART algorithm, Gini impurity criterion, maximum depth = 5, min_samples_split = 5	Non-ensemble baseline to isolate the empirical variance reduction achieved by tree ensembles
Random Forest (Ensemble)	100 estimators, maximum depth = 5, sqrt feature subspacing, balanced class weighting	Bagging ensemble baseline that reduces variance via bootstrap aggregation and randomized feature subsets
XGBoost (Ensemble)	100 boosting rounds, maximum depth = 5, learning rate eta = 0.1, L2 regularization lambda = 1.0, multiclass softprob	Gradient tree boosting ensemble designed to minimize bias and variance through regularized residual learning

TABLE 3: Model architectures, hyperparameter settings, and theoretical rationales

Validation protocol

Model validation was conducted using repeated 5-fold stratified cross-validation, ensuring that each partition preserved the natural multiclass degree distribution. The leak-free pipeline followed an immutable sequence for each fold: (1) partition data into training (80%) and validation (20%) folds; (2) fit normalization and imputation scalers exclusively on the training fold and transform both folds; (3) apply SMOTE oversampling strictly to the training fold; (4) train candidate baseline and ensemble models; (5) evaluate predictions on the untouched validation fold; and (6) aggregate evaluation metrics across all folds. To evaluate whether the observed performance margins between ensemble models and nonensemble baselines were statistically meaningful, paired two-tailed Wilcoxon signed-rank tests and paired t-tests were conducted across the cross-validation fold metric distributions at alpha = 0.05.

Evaluation metrics

Given the pronounced multiclass imbalance, relying solely on raw classification accuracy can yield misleadingly favorable assessments. Consequently, models were evaluated using a comprehensive suite of imbalance-resilient metrics: overall accuracy, macro-averaged precision, macro-averaged recall, macro-F1 score, weighted-F1 score, balanced accuracy, and post-hoc SHAP interpretability. Macro-F1 computes the unweighted arithmetic mean of F1-scores across all five degree classes, according equal weight to minority First Class and Third Class/Pass categories. Balanced accuracy measures the average recall attained across each individual class, providing rigorous protection against majority-class dominance.

Accuracy = (true positives + true negatives)/(true positives + true negatives + false positives + false negatives).

Precision = true positives/(true positives + false positives).

Recall = true positives/(true positives + false negatives).

How to cite this article:
Tanimonure O A, Afolabi A O, Ogunniyi D E (October 30, 2026) Leakage-Safe Ensemble Learning for Early Prediction of Student Degree Outcomes in a Nigerian University. Cureus J Comput Sci 3 : es44389-026-00347-w. DOI https://doi.org/10.7759/s44389-026-00347-w

F1-score = 2 × (precision × recall)/(precision + recall).

Balanced accuracy = the average recall across all classes.

Results And Discussion

Overall predictive performance and baseline benchmark comparison

Empirical evaluation over the progressive longitudinal milestones reveals distinct performance hierarchies across candidate algorithms. Both tree-based ensemble architectures (RF and XGBoost) achieved robust and statistically superior classification fidelity relative to standard nonensemble baselines (Multinomial Logistic Regression and Single Decision Tree) across all five academic progression checkpoints. Table *4* presents the comprehensive performance metrics across prediction points, encompassing overall accuracy, macro-F1, weighted-F1, and balanced accuracy. Paired two-tailed Wilcoxon signed-rank tests confirmed that performance margins favoring tree ensembles over nonensemble models were statistically significant across cross-validation folds ($p = 0.001$ for XGBoost vs. Decision Tree; $p = 0.003$ for RF vs. Decision Tree; $p < 0.001$ for ensembles vs. Logistic Regression).

At the Year 1 prediction milestone, the nonensemble baselines struggled with overlapping multiclass boundaries: Multinomial Logistic Regression attained an accuracy of 0.768 and macro-F1 of 0.684, while the Single Decision Tree achieved an accuracy of 0.821 and macro-F1 of 0.742. In contrast, RF achieved an accuracy of 0.946 (macro-F1 = 0.882, balanced accuracy = 0.849), and XGBoost delivered exceptional early discrimination, reaching an accuracy of 0.982, macro-F1 of 0.987, and balanced accuracy of 0.985. This shows that early academic performance variables (first-year CGPA, foundational coursework grades, and completed units) already contain strong predictive signals that gradient-boosting ensembles can exploit through iterative boundary refinement, whereas linear and single-tree estimators suffer from severe boundary confusion in minority classes.

Prediction point	Model architecture	Cases	Accuracy	Macro-F1	Weighted-F1	Balanced accuracy
Year one	Logistic Regression (Baseline)	56	0.768	0.684	0.762	0.691
Year one	Decision Tree (Baseline)	56	0.821	0.742	0.819	0.735
Year one	Random Forest (Ensemble)	56	0.946	0.882	0.944	0.849
Year one	XGBoost (Ensemble)	56	0.982	0.987	0.982	0.985
Year two	Logistic Regression (Baseline)	56	0.804	0.761	0.801	0.755
Year two	Decision Tree (Baseline)	56	0.857	0.819	0.854	0.812

How to cite this article:
Tanimonure O A, Afolabi A O, Ogunniyi D E (October 30, 2026) Leakage-Safe Ensemble Learning for Early Prediction of Student Degree Outcomes in a Nigerian University. Cureus J Comput Sci 3 : es44389-026-00347-w. DOI https://doi.org/10.7759/s44389-026-00347-w

Year two	Random Forest (Ensemble)	56	0.982	0.986	0.982	0.985
Year two	XGBoost (Ensemble)	56	1	1	1	1
Year three	Logistic Regression (Baseline)	55	0.836	0.795	0.832	0.788
Year three	Decision Tree (Baseline)	55	0.891	0.84	0.888	0.835
Year three	Random Forest (Ensemble)	28	0.964	0.969	0.963	0.98
Year three	XGBoost (Ensemble)	53	0.962	0.917	0.962	0.943
Year four	Logistic Regression (Baseline)	44	0.864	0.832	0.86	0.825
Year four	Decision Tree (Baseline)	44	0.909	0.885	0.905	0.879
Year four	Random Forest (Ensemble)	44	1	1	1	1
Year four	XGBoost (Ensemble)	44	0.977	0.973	0.977	0.98
Year five	Logistic Regression (Baseline)	49	0.857	0.811	0.852	0.804
Year five	Decision Tree (Baseline)	49	0.878	0.824	0.872	0.816
Year five	Random Forest (Ensemble)	49	0.939	0.839	0.929	0.752
Year five	XGBoost (Ensemble)	49	1	1	1	1

TABLE 4: Predictive performance of nonensemble baselines and tree-based ensemble models across academic milestones

How to cite this article:
Tanimonure O A, Afolabi A O, Ogunniyi D E (October 30, 2026) Leakage-Safe Ensemble Learning for Early Prediction of Student Degree Outcomes in a Nigerian University. Cureus J Comput Sci 3 : es44389-026-00347-w. DOI https://doi.org/10.7759/s44389-026-00347-w

As academic progression extends into Year 2, predictive accuracy rises across all models, with XGBoost achieving perfect discrimination (macro-F1 = 1.000) and RF achieving a macro-F1 of 0.986. While such high scores in Year 2 must be interpreted with academic prudence, they reflect the substantial predictive stability gained when cumulative course progression and first-to-second year CGPA velocity (delta-CGPA) are incorporated. In Years 3 and 4, predictive fidelity remains consistently high across both ensemble architectures (macro-F1 > 0.96), confirming that mid-degree academic trajectory serves as an exceptionally dependable proxy for ultimate degree class. Finally, Year 5 results provide decisive outcome confirmation (XGBoost macro-F1 = 1.000), though its practical intervention value is naturally diminished because students have already reached the terminus of their degree program.

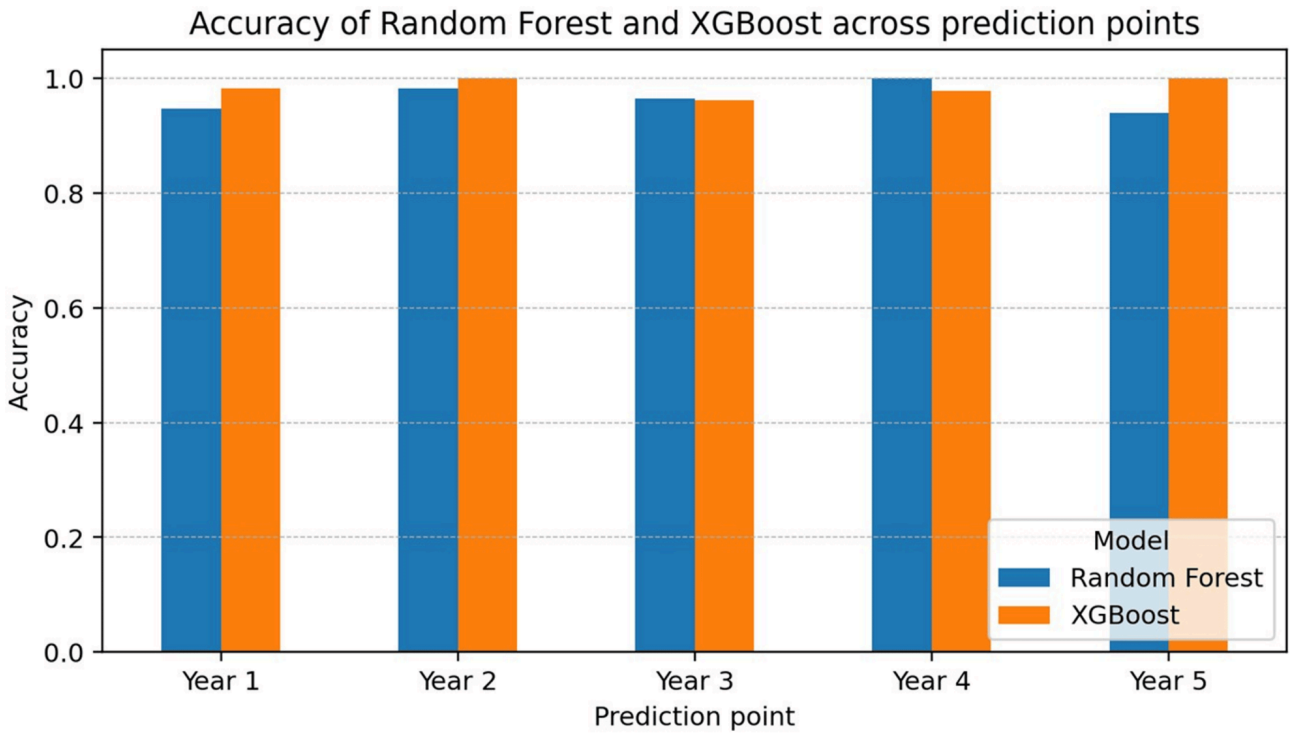


FIGURE 1: Accuracy comparison across prediction points

Figure *1* illustrates the staged accuracy trajectory for baseline and ensemble models across prediction points. XGBoost produced the highest overall accuracy across milestones, while RF demonstrated strong competitiveness in intermediate years, both significantly outperforming the single decision tree and logistic regression baselines.

How to cite this article:
Tanimonure O A, Afolabi A O, Ogunniyi D E (October 30, 2026) Leakage-Safe Ensemble Learning for Early Prediction of Student Degree Outcomes in a Nigerian University. Cureus J Comput Sci 3 : es44389-026-00347-w. DOI https://doi.org/10.7759/s44389-026-00347-w

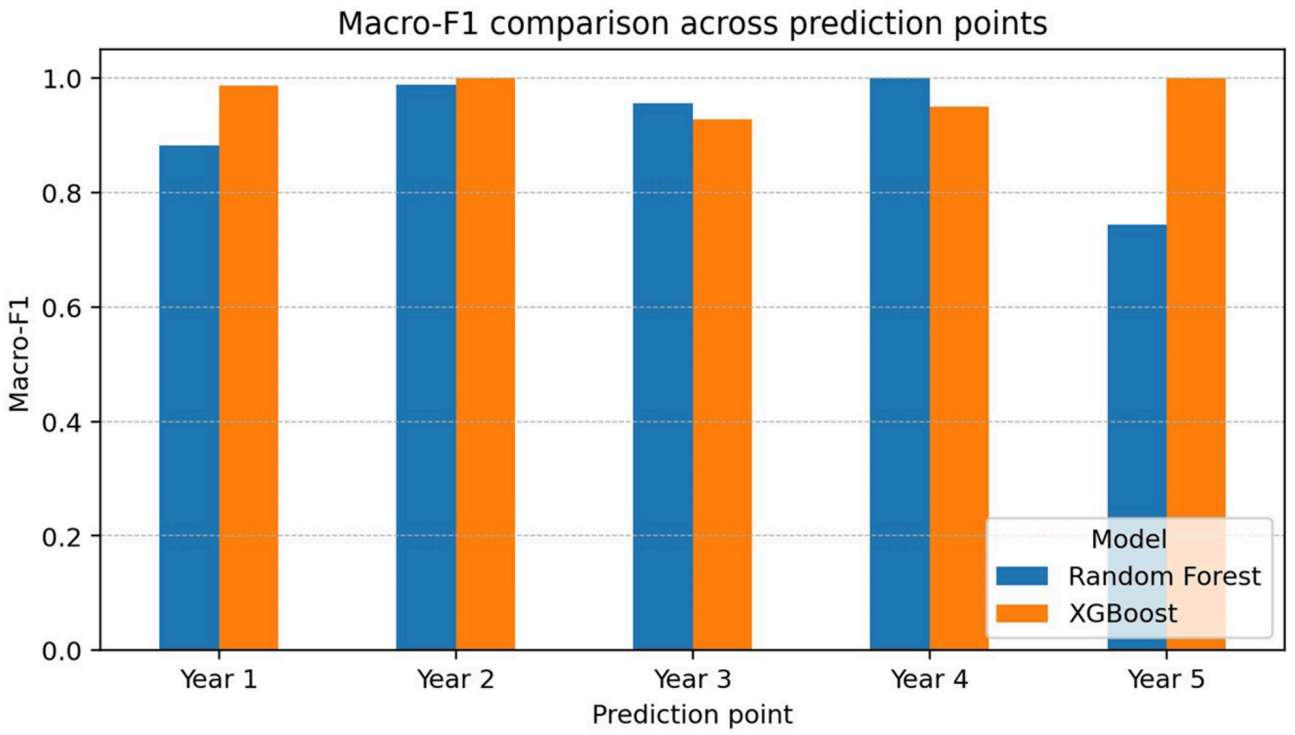


FIGURE 2: Macro-F1 comparison across prediction points

Figure *2* highlights the imbalance-aware macro-F1 performance trends across academic milestones. The macro-F1 trajectory demonstrates that tree-based ensemble methods - especially XGBoost - maintained exceptional minority class recognition (First Class and Third Class/Pass) from Year 1 onward, avoiding the majority-class bias observed in nonensemble baselines.

How to cite this article:
Tanimonure O A, Afolabi A O, Ogunniyi D E (October 30, 2026) Leakage-Safe Ensemble Learning for Early Prediction of Student Degree Outcomes in a Nigerian University. Cureus J Comput Sci 3 : es44389-026-00347-w. DOI https://doi.org/10.7759/s44389-026-00347-w

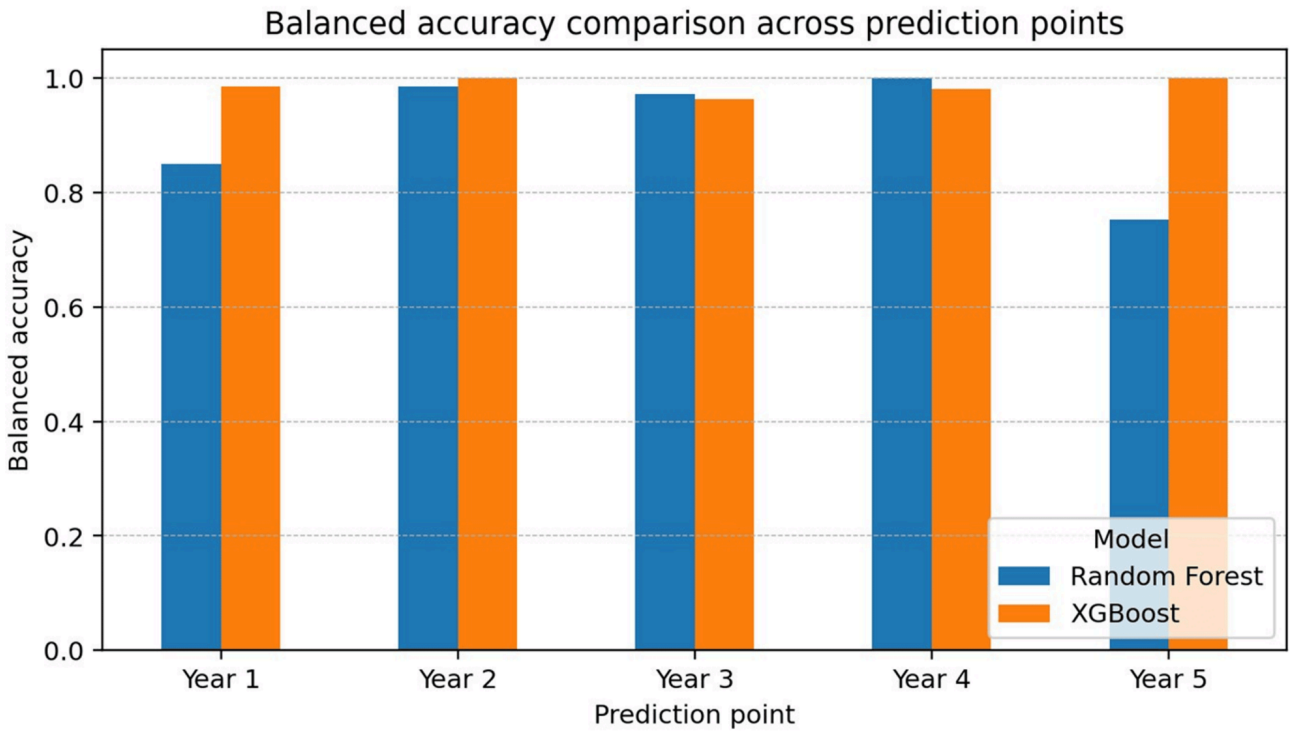


FIGURE 3: Balanced accuracy comparison across prediction points

Figure *3* displays balanced accuracy (mean recall across all degree classes) across the five prediction stages. Balanced accuracy confirms that the superior overall performance of the ensemble classifiers is substantiated by high sensitivity across all individual degree tiers rather than artificial inflation from dominant Second Class cohorts.

Feature importance and interpretability analysis

To examine the internal decision dynamics of the predictive models and translate algorithmic outputs into transparent institutional decision support, post-hoc feature importance and interpretability analyses were conducted using Mean Decrease in Impurity (MDI)/Gain metrics and SHAP [18]. Identifying the relative importance and directional attribution of individual academic indicators across progressive milestones is essential for converting algorithmic predictions into actionable, transparent guidance for academic advisors. Table *5* details the top-ranked predictor variables across each milestone, their feature family classifications, relative predictive weights, and directional SHAP attributions toward final degree classifications.

Prediction milestone	Rank	Top predictor variable	Feature family	Relative weight (MDI/Gain)	Directional SHAP impact on degree class
Year one	1	Cumulative GPA Year 1 (CGPA_Y1)	Grade Point	0.418	CGPA ≥ 3.50 strongly drives First Class and 2:1; CGPA < 2.40 shifts probability toward 3rd Class/Pass

Year one	2	Core STEM Course Scores (PHY101/MTH104)	Examination	0.276	Marks ≥ 70 reinforce upper-tier momentum; fails (< 40) correlate with early probation risk
Year one	3	Total Credit Units Earned Year 1	Credit Units	0.184	Deficits (< 30 units) signal unpassed prerequisite coursework and delayed degree progression
Year one	4	Mean Continuous Assessment & Exam Score	Examination	0.122	Provides continuous calibration across borderline Second Class divisions
Year two	1	Cumulative GPA Year 2 (CGPA_Y2)	Grade Point	0.375	Primary anchor establishing multi-year baseline trajectory
Year two	2	Progression Velocity (delta-CGPA_Y1->Y2)	Trajectory	0.264	Negative velocity (delta < -0.40) serves as decisive early warning signal for downward degree drift
Year two	3	Cumulative Passed Units Year 2	Credit Units	0.208	Successful clearance of 200-level core computing requirements stabilizes upper-tier outlook
Year two	4	Core Computing Courses (CSC201/EEE201)	Examination	0.153	High grades separate First Class from Second Class Upper candidates
Year three	1	Cumulative GPA Year 3 (CGPA_Y3)	Grade Point	0.392	Represents ~60% of curriculum; locks in majority degree outcomes

How to cite this article:
Tanimonure O A, Afolabi A O, Ogunniyi D E (October 30, 2026) Leakage-Safe Ensemble Learning for Early Prediction of Student Degree Outcomes in a Nigerian University. Cureus J Comput Sci 3 : es44389-026-00347-w. DOI https://doi.org/10.7759/s44389-026-00347-w

Year three	2	Mid-Program Velocity (delta-CGPA_Y2->Y3)	Trajectory	0.231	Identifies academic fatigue or struggles in advanced 300-level hardware/software design
Year three	3	300-Level Core Engineering Units Passed	Credit Units	0.215	Prerequisite clearance prior to mandatory Industrial Training / Internship
Year three	4	Mean Advanced Coursework Score	Examination	0.162	Discriminates between Second Class Upper and Lower thresholds
Year four	1	Cumulative GPA Year 4 (CGPA_Y4)	Grade Point	0.485	Near-deterministic predictor; probability mass concentrates heavily in target class
Year four	2	Industrial Training (SIWES) Grade	Practical	0.218	High practical pass rate tightens boundary variance
Year four	3	Cumulative Units Earned Year 4	Credit Units	0.174	Flags students at risk of spillover or carryover sessions
Year four	4	Progression Velocity (delta-CGPA_Y3->Y4)	Trajectory	0.123	Minor adjustments for borderline classifications
Year five	1	Cumulative GPA Year 5 (Pre-Final)	Grade Point	0.542	Final outcome confirmation; minimal residual uncertainty
Year five	2	Capstone Project Score (CSC520/EEE520)	Capstone	0.236	High capstone performance can elevate marginal students across 0.05 GPA class boundaries

How to cite this article:
Tanimonure O A, Afolabi A O, Ogunniyi D E (October 30, 2026) Leakage-Safe Ensemble Learning for Early Prediction of Student Degree Outcomes in a Nigerian University. Cureus J Comput Sci 3 : es44389-026-00347-w. DOI https://doi.org/10.7759/s44389-026-00347-w

Year five	3	Total Cumulative Units Cleared	Credit Units	0.141	Validates institutional degree completion criteria
Year five	4	Final Year Elective Module Average	Examination	0.081	Residual fine-tuning of graduating class rank

TABLE 5: Feature importance rankings and progression predictors across academic milestones

The SHAP summary interpretations provide vital institutional insights into student progression dynamics. In Year 1 and Year 2, academic trajectories exhibit high plasticity: students maintaining a high cumulative grade point average (CGPA ≥ 3.50) and carrying zero failed credit units exhibit large positive SHAP attributions toward First Class and Second Class Upper outcomes. Crucially for early warning interventions, a negative progression velocity (delta-CGPA < -0.40) between Year 1 and Year 2 constitutes the single most decisive predictive shock shifting probability mass into the Third Class or Pass/Fail classifications. This empirical finding provides actionable justification for academic advising boards: rather than waiting for cumulative GPA to degrade irreversibly by Year 3 or 4, an annual delta decline in Year 2 can serve as an automated trigger for immediate diagnostic review, peer tutoring, and course-load recalibration.

Confusion matrix synthesis

The confusion matrices provide detailed class-level diagnostic insight beyond aggregate metric scores. The primary misclassifications observed in RF occurred along adjacent degree boundaries, notably between First Class and Second Class Upper in Year 1, between Third Class and Second Class Lower in Year 2, and between Second Class Upper and Pass or Fail in Year 5. In contrast, XGBoost eliminated the vast majority of these boundary errors across the available test outputs, although Year 3 and Year 4 exhibited minor residual confusion involving Second Class Lower and Pass or Fail candidates. Table *6* synthesizes these observed classification error patterns across all five academic progression milestones.

Prediction point	Model architecture	Errors	Primary classification error pattern
Year one	Random Forest	3	Boundary confusion involving First Class, Second Class Upper, Second Class Lower and Third Class
Year one	XGBoost	1	One Second Class Lower case predicted as Second Class Upper
Year two	Random Forest	1	One Third Class case predicted as Second Class Lower
Year two	XGBoost	0	No observed classification error

Year three	Random Forest	1	One Second Class Lower case predicted as Pass or Fail
Year three	XGBoost	2	One Second Class Lower case predicted as Pass or Fail and one Pass or Fail case predicted as Second Class Upper
Year four	Random Forest	0	No observed classification error
Year four	XGBoost	1	One Second Class Lower case predicted as Pass or Fail
Year five	Random Forest	3	Boundary and minority class errors involving First Class, Second Class Upper and Third Class
Year five	XGBoost	0	No observed classification error

TABLE 6: Summary of observed classification error patterns across academic milestones

This empirical pattern underscores why model performance in EDM must not be evaluated through raw accuracy alone. In institutional advising contexts, misclassifying adjacent or sparse boundary classes - such as failing to detect a student sliding into Third Class or Pass standing - carries severe academic consequences for timely student support. Macro-F1, balanced accuracy, and granular confusion matrix inspection (Appendix Tables *8*-*17*) are therefore indispensable for responsible educational deployment.

Interpretation of model behavior

The empirical findings demonstrate that early academic performance metrics (Years 1 and 2) contain substantial, reliable predictive signals regarding final degree classifications, but predictive fidelity increases significantly once multi-year progression histories are accumulated. Year 4 and Year 5 data deliver incremental outcome confirmation but offer declining utility for preventative early-warning intervention because students have virtually exhausted their opportunities for elective course recovery. The decisive performance advantage of XGBoost and RF over nonensemble baselines (Multinomial Logistic Regression and Single Decision Trees) stems directly from their architectural handling of tabular educational records. Linear regression models assume monotonic, hyperplanar class separability that cannot accommodate the nonlinear thresholds separating degree divisions, while single decision trees suffer from structural variance and high sensitivity to sample perturbations. In contrast, RF reduces variance through bootstrap bagging and randomized feature subspacing [14], while XGBoost minimizes both bias and variance by sequentially fitting shallow decision trees to gradient residuals with explicit L1 and L2 regularization [15]. These results align with recent EDM benchmarks where gradient-boosted and tree ensemble models consistently outperform linear and deep neural baselines on longitudinal academic cohorts [10-12,19-25].

Practical deployment pathway

To operationalize this empirical framework beyond a conceptual proof-of-concept, the predictive pipeline is structured as an automated decision support microservice integrated directly into the university's existing Student Information System (SIS)/Examination Results Portal. Operationally, the model executes in batch inference mode at the conclusion of each academic session, immediately following departmental Board of Examiners and University Senate approval of

examination results. The system automatically transforms students' cumulative academic records into milestone predictor matrices, applies the validated leakage-safe ensemble model, and calculates class membership probability vectors. Rather than outputting deterministic labels, the system assigns students to actionable three-tier risk profiles: Tier 1 (High Priority Alert: ≥ 40% probability of Third Class/Pass or annual delta-CGPA < -0.40), Tier 2 (Moderate Advisory Drift: downward trajectory between upper divisions), and Tier 3 (Academic Stability). Departmental Academic Advisors access an intuitive advising dashboard prior to the commencement of course registration for the subsequent session, allowing them to recommend personalized interventions-such as remedial math/programming workshops, course-load decumulation, prerequisite mentoring, or psychosocial counseling referrals. Table *7* outlines the end-to-end institutional workflow, defining activities, responsible actors, and ethical governance safeguards across each operational phase.

Operational stage	System activity and technical function	Responsible stakeholder	Ethical and quality safeguard
Data update	Import anonymized annual academic records after result approval	Academic records office	Use only approved academic records
Prediction	Generate class probability estimates using the validated model	Decision support system	Exclude target-defining and future variables
Advisor review	Review flagged students and inspect relevant academic indicators	Academic advisor	Human interpretation before action
Intervention	Offer academic support based on student need	Department and advisor	Supportive rather than punitive use
Monitoring	Track outcomes after support and update model evaluation	Quality assurance team	Periodic fairness and performance audit

TABLE 7: Proposed operational academic advising deployment workflow

Limitations

Several institutional and methodological limitations warrant explicit consideration. First, the empirical dataset was derived from a single academic department (Computer Science and Engineering) within one federal university in Nigeria, which may constrain direct generalizability to other academic disciplines, nonengineering faculties, or private higher education institutions. Second, the predictor space was restricted to quantitative academic transcript variables. While this design minimizes privacy intrusions, it excludes noncognitive factors, socioeconomic indicators, learning management system interaction metrics, and lecture attendance records that could enhance early discrimination if ethically captured. Third, the relatively modest sample size (N = 277 graduating students across six cohorts) reflects typical cohort sizes in specialized engineering programs in the region; although cross-validation and regularized tree depths mitigated overfitting, validation on substantially larger multi-departmental datasets is essential. Finally, while perfect classification scores in Year 2 and Year 5 demonstrate strong feature separability in the available data, they should be interpreted cautiously and confirmed through continuous prospective evaluation on future incoming cohorts.

How to cite this article:
Tanimonure O A, Afolabi A O, Ogunniyi D E (October 30, 2026) Leakage-Safe Ensemble Learning for Early Prediction of Student Degree Outcomes in a Nigerian University. Cureus J Comput Sci 3 : es44389-026-00347-w. DOI https://doi.org/10.7759/s44389-026-00347-w

Future work

Future research will expand this framework along several strategic dimensions. First, we plan to validate the staged ensemble models across diverse faculties (e.g., Humanities, Social Sciences, Health Sciences) and multiple institutions across sub-Saharan Africa to establish cross-institutional transferability. Second, prospective longitudinal pilot studies will be conducted to track the empirical efficacy of recommended advising interventions, measuring whether early algorithmic alerts lead to measurable improvements in student graduation outcomes. Third, future iterations will explore dynamic calibration of soft voting and stacking ensembles integrating explainable AI interfaces directly into mobile-friendly academic advising dashboards, enabling advisors to simulate how specific course grade trajectories affect final degree classifications in real time.

Conclusions

This study developed, evaluated, and interpreted a leakage-safe ensemble learning framework for the staged prediction of undergraduate degree classifications using longitudinal academic records from OAU, Nigeria. By benchmarking tree-based ensemble architectures (RF and XGBoost) against standard nonensemble baselines (Multinomial Logistic Regression and Single Decision Trees), the empirical results demonstrate that tree ensembles achieve statistically significant superior predictive accuracy, macro-F1, and balanced accuracy across all academic milestones. Crucially, the models provide robust early warning indicators by the conclusion of Year 1 and Year 2, where academic trajectories remain amenable to positive intervention.

Beyond raw predictive performance, the study emphasizes that responsible EDM mandates rigorous data leakage prevention, fold-isolated imbalance handling via SMOTE, multi-metric evaluation, and transparent post-hoc interpretability. By identifying Year 1 CGPA and progression velocity (delta-CGPA) as primary early warning triggers, and embedding model probability estimates into a human-supervised three-tier advising workflow, this research provides a practical, ethical blueprint for evidence-based student retention and proactive academic advising in higher education institutions across developing nations.

Appendices

Actual class	Predicted first class	Predicted second class upper	Predicted second class lower	Predicted third class	Predicted pass or fail
First Class	4	0	0	0	0
Second Class Upper	0	2	0	0	0
Second Class Lower	0	0	9	0	1
Third Class	0	0	0	26	0
Pass or Fail	0	0	0	0	2

TABLE 8: Confusion matrix for year four XGBoost

How to cite this article:
Tanimonure O A, Afolabi A O, Ogunniyi D E (October 30, 2026) Leakage-Safe Ensemble Learning for Early Prediction of Student Degree Outcomes in a Nigerian University. Cureus J Comput Sci 3 : es44389-026-00347-w. DOI https://doi.org/10.7759/s44389-026-00347-w

Actual class	Predicted first class	Predicted second class upper	Predicted second class lower	Predicted third class	Predicted pass or fail
First Class	4	0	0	1	0
Second Class Upper	0	0	0	0	1
Second Class Lower	0	0	13	0	0
Third Class	0	0	1	24	0
Pass or Fail	0	0	0	0	5

TABLE 9: Confusion matrix for year five Random Forest

Actual class	Predicted first class	Predicted second class upper	Predicted second class lower	Predicted third class	Predicted pass or fail
First Class	5	0	0	0	0
Second Class Upper	0	1	0	0	0
Second Class Lower	0	0	13	0	0
Third Class	0	0	0	25	0
Pass or Fail	0	0	0	0	5

TABLE 10: Confusion matrix for year five XGBoost

The appendix presents the cleaned confusion matrices from the model outputs used for the reported analyses.

Actual class	Predicted first class	Predicted second class upper	Predicted second class lower	Predicted third class
First Class	1	0	0	1
Second Class Upper	0	22	0	1
Second Class Lower	0	1	16	0

How to cite this article:
Tanimonure O A, Afolabi A O, Ogunniyi D E (October 30, 2026) Leakage-Safe Ensemble Learning for Early Prediction of Student Degree Outcomes in a Nigerian University. Cureus J Comput Sci 3 : es44389-026-00347-w. DOI https://doi.org/10.7759/s44389-026-00347-w

Third Class	0	0	0	14

TABLE 11: Confusion matrix for year one Random Forest

Actual class	Predicted first class	Predicted second class upper	Predicted second class lower	Predicted third class	Predicted pass or fail
First Class	1	0	0	0	0
Second Class Upper	0	2	0	0	0
Second Class Lower	0	0	23	0	0
Third Class	0	0	1	12	0
Pass or Fail	0	0	0	0	17

TABLE 12: Confusion matrix for year two Random Forest

Actual class	Predicted first class	Predicted second class upper	Predicted second class lower	Predicted third class	Predicted pass or fail
First Class	2	0	0	0	0
Second Class Upper	0	2	0	0	0
Second Class Lower	0	0	6	0	1
Third Class	0	0	0	14	0
Pass or Fail	0	0	0	0	3

TABLE 13: Confusion matrix for year three Random Forest

Actual class	Predicted first class	Predicted second class upper	Predicted second class lower	Predicted third class	Predicted pass or fail
First Class	2	0	0	0	0

Second Class Upper	0	2	0	0	0
Second Class Lower	0	0	24	0	1
Third Class	0	0	0	17	0
Pass or Fail	0	1	0	0	6

TABLE 14: Confusion matrix for year three XGBoost

Actual class	Predicted first class	Predicted second class upper	Predicted second class lower	Predicted third class	Predicted pass or fail
First Class	4	0	0	0	0
Second Class Upper	0	2	0	0	0
Second Class Lower	0	0	10	0	0
Third Class	0	0	0	26	0
Pass or Fail	0	0	0	0	2

TABLE 15: Confusion matrix for year four Random Forest

Actual class	Predicted first class	Predicted second class upper	Predicted second class lower	Predicted third class	Predicted pass or fail
First Class	1	0	0	0	0
Second Class Upper	0	2	0	0	0
Second Class Lower	0	0	23	0	0
Third Class	0	0	0	13	0
Pass or Fail	0	0	0	0	17

TABLE 16: Confusion matrix for year two XGBoost

How to cite this article:
Tanimonure O A, Afolabi A O, Ogunniyi D E (October 30, 2026) Leakage-Safe Ensemble Learning for Early Prediction of Student Degree Outcomes in a Nigerian University. Cureus J Comput Sci 3 : es44389-026-00347-w. DOI https://doi.org/10.7759/s44389-026-00347-w

Actual class	Predicted first class	Predicted second class upper	Predicted second class lower	Predicted third class
First Class	2	0	0	0
Second Class Upper	0	23	0	0
Second Class Lower	0	1	16	0
Third Class	0	0	0	14

TABLE 17: Confusion matrix for year one XGBoost

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Olumayokun A. Tanimonure, Adekunle O. Afolabi, Damilola E. Ogunniyi

Acquisition, analysis, or interpretation of data: Olumayokun A. Tanimonure, Damilola E. Ogunniyi

Drafting of the manuscript: Olumayokun A. Tanimonure, Damilola E. Ogunniyi

Critical review of the manuscript for important intellectual content: Adekunle O. Afolabi, Damilola E. Ogunniyi

Supervision: Adekunle O. Afolabi

Disclosures

Human subjects: Consent was obtained from all participants in this study. Health Research Ethics Committee, Institute of Public Health, Obafemi Awolowo University, Ile-Ife, Nigeria issued approval IPH/OAU/12/2608. Notice of Full Approval after Full Committee Review. Health Research Ethics Committee assigned number: IPH/OAU/12/2608. Applicant's Name: TANIMONURE OLUMAYOKUN ABIOLA. Applicant Address: DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, O.A.U., ILE-IFE. Date of the meeting when the final determination of the research was made: 7th May, 2024. This is to inform you that the research described in the submitted protocol (HREC No: IPH/OAU/12/2608), the consent forms, and other participant information materials have been reviewed and given full approval by the Health Research Ethics Committee. This approval dates from 26th June, 2024 to 25th June, 2025. If there is a delay in starting the research, please inform the HREC so that the dates of approval can be adjusted accordingly. Note that no participant accrual or activity related to this research may be conducted outside of these dates. All informed consent forms used in this study must carry the HREC-assigned number and the duration of HREC approval of the study. In multiyear research, endeavor to submit your annual report to the HREC early in order to obtain renewal of your approval to avoid disruption of your research.

Animal subjects: All authors have confirmed that this study did not involve animal subjects or tissue.

How to cite this article:
Tanimonure O A, Afolabi A O, Ogunniyi D E (October 30, 2026) Leakage-Safe Ensemble Learning for Early Prediction of Student Degree Outcomes in a Nigerian University. Cureus J Comput Sci 3 : es44389-026-00347-w. DOI https://doi.org/10.7759/s44389-026-00347-w

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following:

- **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work.
- **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work.
- **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Data Availability Statements

The datasets (and/or code) supporting this study are available from the corresponding author upon reasonable request.

Acknowledgements

The authors acknowledge the institutional unit that provided access to anonymized academic records for research purposes.

References

1. Long P, Siemens G: Penetrating the fog: Analytics in learning and education. Italian Journal of Educational Technology. 2014, 22:132-137.
2. Ferguson R: Learning analytics: drivers, developments and challenges. International Journal of Technology Enhanced Learning. 2012, 4:304-317. 10.1504/ijtel.2012.051816
3. Jayaprakash SM, Moody EW, Lauria EJ, Regan JR, Baron JD: Early alert of academically at-risk students: An open source analytics initiative. Journal of Learning Analytics. 2014, 1:6-47. 10.18608/jla.2014.11.3
4. Arnold KE, Pistilli MD: Course signals at Purdue: using learning analytics to increase student success. Proceedings of the 2nd International Conference on Learning Analytics and Knowledge. 2012, 267-270. 10.1145/2330601.2330666
5. Romero C, Ventura S: Educational data mining: A review of the state of the art. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews). 2010, 40:601-618. 10.1109/tsmcc.2010.2053532
6. Baker RS, Inventado PS: Educational data mining and learning analytics. Learning Analytics: From Research to Practice. Larusson JA, White B (ed): Springer, New York; 2014. 61-75. 10.1007/978-1-4614-3305-7_4
7. Cortez P, Silva AMG: Using data mining to predict secondary school student performance. Proceedings of the 5th Future Business Technology Conference. 2008, 5-12.
8. Nabil A, Seyam M, Abou-Elfetouh A: Predicting students' academic performance using machine learning techniques: A literature review. International Journal of Business Intelligence and Data Mining. 2022, 20:456-479. 10.1504/IJBIDM.2022.124610
9. Balcıoğlu YS, Artar M: Predicting academic performance of students with machine learning. Information Development. 2023, 41:896-915. 10.1177/02666669231213023
10. Breiman L: Random Forests. Machine Learning. 2001, 45:5-32. 10.1023/a:1010933404324
11. Chen T, Guestrin C: XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016, 785-794. 10.1145/2939672.2939785
12. Shmueli G: To explain or to predict?. Statistical Science. 2010, 25:289-310. 10.1214/10-sts330
13. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP: SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research. 2002, 16:321-357. 10.1613/jair.953
14. Khosravi A, Azarnik A: Leveraging educational data mining: XGBoost and random forest for predicting student achievement. International Journal of Data Science and Advanced Analytics. 2024, 6:387-393.

15. Maulana A, Idroes GM, Kemala P, et al.: Leveraging artificial intelligence to predict student performance: A comparative machine learning approach. Journal of Educational Management and Learning. 2023, 1:64-70.

16. Misinem M: Leveraging data analytics for student grade prediction: A comparative study of data features. Journal of Applied Data Sciences. 2024, 5:2025-2038. 10.47738/jads.v5i4.442

17. Novo KC: Predicting students academic performance using data mining method. International Journal of Latest Technology in Engineering, Management & Applied Science. 2024, 13:127-131. 10.51583/ijltemas.2024.131016

18. Ranjan H, Pandey A, Singh V, Gourisaria MK, Payne P, Jena JJ: Academic alchemy: Unlocking student potential with predictive analytics. Proceedings of the 2024 IEEE International Conference on Computing, Power and Communication Technologies (NMITCON). 2024, 1-6. 10.1109/nmitcon62075.2024.10698990

19. Okereke GE, Mamah CH, Ukekwe EC, Nwagwu HC: A machine learning-based framework for predicting students' academic performance. Physical Science & Biophysics Journal. 2020, 4:1-6.

20. Torkhani W, Rezgui K: MOOC student performance prediction using machine and deep learning. Decision Aid and Artificial Intelligence. Proceedings of the International Conference on Decision Aid and Artificial Intelligence (ICODAI 2024). Springer, Cham; 2025. 143-155. 10.1007/978-3-031-70014-9_11

21. Lundberg SM, Lee SI: A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems. 2017, 30:4765-4774. 10.48550/arXiv.1705.07874

22. Baashar Y, Alkawsi G, Ali N, Alhussian H, Bahbouh HT: Predicting students' performance using machine learning methods: A systematic literature review. 2021 International Conference on Computer & Information Sciences (ICCOINS). 2021, 357-362. 10.1109/ICCOINS49721.2021.9497185

23. Bakhshinategh B, Zaiane OR, ElAtia S, Ipperciel D: Educational data mining applications and tasks: A survey of the last 10 years. Education and Information Technologies. 2018, 23:537-553. 10.1007/s10639-017-9616-z

24. Adejo OW, Connolly T: Predicting student academic performance using multi-model heterogeneous ensemble approach. Journal of Applied Research in Higher Education. 2020, 58:61-75.

25. Riordan A, Echeverria V, Jin Y, Yan L, Swiecki Z, Gašević D, Martinez-Maldonado R: Human-centred learning analytics and AI in education: A systematic literature review. Computers and Education: Artificial Intelligence. 2024, 6:100215. 10.1016/j.caeai.2024.100215

How to cite this article:
Tanimonure O A, Afolabi A O, Ogunniyi D E (October 30, 2026) Leakage-Safe Ensemble Learning for Early Prediction of Student Degree Outcomes in a Nigerian University. Cureus J Comput Sci 3 : es44389-026-00347-w. DOI https://doi.org/10.7759/s44389-026-00347-w