Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Machine Learning Prediction of Undergraduate Academic Performance for Early Identification of At-Risk Students: Evidence from Nnamdi Azikiwe University

Chinenye Amaonye

cb.amaonye@unizik.edu.ng

Nnamdi Azikiwe University https://orcid.org/0000-0001-9145-4339

Judith Nneka OKAFOR

Nnamdi Azikiwe University https://orcid.org/0009-0006-0911-8998

Samuel Oweh Abang

abangsamueloweh@gmail.com

University of Calabar https://orcid.org/0000-0002-8579-238X

Ngozi Eucharia Anyikwa

Nnamdi Azikiwe University https://orcid.org/0009-0001-6539-6697

Juliana Uzonna UMEOZOR

Nnamdi Azikiwe University

Research Article

Keywords: Machine Learning, Prediction, At-Risk-Students, Academic Performance, Early Identification,

Posted Date: October 28th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11516202/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Machine Learning Prediction of Undergraduate Academic Performance for Early Identification of At-Risk Students: Evidence from Nnamdi Azikiwe University

Chinenye Blessing AMAONYE[1] cb.amaonye@unizik.edu.ng https://orcid.org/0000-0001-9145-4339 +2348034118503

Judith Nneka OKAFOR[1] ju.okafor@unizik.edu,ng https://orcid.org/0009-0006-0911-8998
Samuel Oweh Abang[2] abangsamueloweh@gmail.com https://orcid.org/0000-0002-8579-238X
Ngozi Eucharia ANYIKWA[3] ne.anyikwa@unizik.edu.ng https://orcid.org/009-0001-6539-6697

Juliana Uzonna UMEOZOR[1] julzyiana2@gmail.com +2348035085408

[1,"]Department of Educational Management and Policy, Faculty of Education, Nnamdi Azikiwe University, Awka, Nigeria
[2]Departmenet of Economics, Faculty of Social Sciences, University of Calabar, Nigeria.
[3]Department of Early Childhood and Primary Education, Faculty of Education, Nnamdi Azikiwe University, Awka, Nigeria

Correspondence: cb.amaonye@unizik.edu.ng

Funding: This research receives no funding.

Availability of data and material: Data for the research can be provided upon reasonable request.

Conflict of Interest: The authors declare no conflict of interest.

Declaration of AI Assisted Technologies: Machine learning models (Logistic Regression, Random Forest, Support Vector Machine (SVM) with radial basis function, and XGBoost) were used in this study They are used in training and analyzing data. No AI tools was used in the writing of the manuscript.

Consent to participate: Not applicable. No human participants used in the study.

Consent to publish: Not applicable.(

Abstract

The increasing availability of student academic records provides an opportunity for universities to identify students who may struggle academically before serious problems develop. This study examined the use of machine learning techniques to predict undergraduate academic performance and support the early identification of at-risk students at Nnamdi Azikiwe University. The study was motivated by the need for timely academic support, particularly in a setting where intervention often comes after students have already experienced poor academic outcomes. Specifically, the study assessed the ability of selected machine learning models to classify students according to academic risk, compared their predictive performance, and identified the variables that contributed most to the predictions. A quantitative predictive design was adopted using records for 400 undergraduate students. The data included admission scores, first-semester GPA, continuous assessment,

attendance, carryover courses, age, gender and mode of entry. Logistic Regression, Random Forest, Support Vector Machine (SVM) with radial basis function, and XGBoost were trained using 10-fold cross-validation and evaluated on an independent test sample using accuracy, precision, recall, F1-score and ROC-AUC. The results showed that 99 students (24.8%) were classified as at risk. First-semester GPA, carryover courses and continuous assessment were the most important predictors. On the test data, SVM recorded the highest accuracy (82.5%), F1-score (69.6%) and ROC-AUC (0.911), while Logistic Regression also produced strong results. The findings indicate that machine learning can provide useful early warning signals from routine university records. The study recommends that universities incorporate predictive early-warning systems into academic advising and student support programmes. Such systems should complement, rather than replace, lecturers' and advisers' professional judgement, with predictions used to trigger timely support rather than permanently label students.

Keywords*: Machine Learning, Prediction, At-Risk-Students, Academic Performance, Early Identification,*

1.0 INTRODUCTION

The expansion of higher education has created an increasingly important challenge: universities now have to support much larger and more diverse student populations while identifying academic difficulties early enough for intervention to matter. Global higher-education enrolment reached about 254 million students in 2022, more than twice the level recorded in 2000, reflecting the rapid expansion and diversification of tertiary education worldwide (UNESCO, 2024). This expansion has also increased the volume of academic, administrative and learning-related data generated by universities. The growing availability of such data has changed how student success can be understood. Rather than waiting until students fail examinations or accumulate poor grades, institutions can increasingly use data-driven methods to identify patterns associated with academic difficulty and provide support before poor performance becomes difficult to reverse.

Machine learning has become particularly relevant to this shift because it can identify complex relationships among multiple student characteristics and use these relationships to classify or predict future outcomes. In higher education, academic performance prediction has developed alongside learning analytics and educational data mining, with applications ranging from predicting grade point average and retention to identifying students who may require additional academic support. Helal et al. (2018), for example, showed that combining enrolment characteristics with learning-management-system activity improved the identification of vulnerable students, while also demonstrating that predictive performance can differ across student groups. Musso et al. (2020) similarly showed that machine-learning approaches can be used to predict important educational outcomes such as academic achievement, retention and degree completion. These developments are important because

prediction is most useful when it leads to timely and targeted intervention rather than simply producing an accurate classification after the academic problem has already occurred.

The value of early prediction is evident in empirical studies. Gray and Perkins (2019) demonstrated that machine-learning models could identify students at risk of failure as early as the third week of a semester, with reported pass-fail prediction accuracy of about 97% in their institutional setting. Akçapınar, Altun and Aşkar (2019), using interaction data from university students, found that unsuccessful students could be identified as early as the third week, although predictive performance varied according to the available features and modelling approach. More recent evidence also suggests that early prediction does not necessarily require highly complex data environments. A 2023 study using learning-management-system logs found that at-risk students could be identified at approximately one-quarter of course completion, although the authors stressed the importance of validation across different institutional contexts.

The relevance of this approach is particularly strong in developing higher-education systems where institutional resources for individual academic monitoring may be limited. In Nigeria, the rapid expansion of university education has increased the need for more systematic approaches to monitoring student progression. Existing Nigerian evidence already shows that machine learning can provide useful predictive information. Ekubo and Esiefarienrhe (2022), using undergraduate records from Niger Delta University, compared five classification algorithms for predicting low academic performance and reported substantial differences in predictive performance across the models. Their findings demonstrate both the potential of machine learning in the Nigerian university environment and the importance of testing alternative algorithms rather than assuming that one model will perform consistently across settings.

The problem, however, is not simply whether academic performance can be predicted. The more important question is whether prediction can be sufficiently early, reliable and context-sensitive to support academic intervention. Recent evidence reinforces this concern. Çırak, Akıllı and Ekinci (2024), using data from 11,698 first-year students, found that random forest models achieved 90.2% accuracy, with previous academic performance, programme and university placement examination scores among important predictors. At the same time, the study showed that model performance depended strongly on the information available to the prediction system. Similarly, research on higher-education achievement has increasingly emphasized interpretable models because prediction without an understanding of influential factors may have limited practical value for academic decision-making.

This issue is also evident within the Nigerian university context. Evidence from Nnamdi Azikiwe University indicates that undergraduate academic performance can be affected by several institutional and student-related conditions, including examination scheduling, financial pressures, lecture demands and other sources of student stress. A study of the university reported a student population of 29,976 at the time of its investigation and identified several conditions associated with student stress and academic performance. Such evidence suggests that academic outcomes are not determined by a single characteristic but emerge from the interaction of academic history, engagement, attendance, study behaviour and institutional circumstances.

Despite the growing literature, an important gap remains. Much of the existing evidence comes from online learning environments, single courses, particular programmes or institutions outside Nigeria. Even within Nigeria, studies have largely focused on individual universities, specific disciplines or admission-related prediction, making it difficult to assume that findings from one institution will transfer directly to another. Recent Nigerian work has also demonstrated that admission variables alone may provide limited information about subsequent academic performance, suggesting the need for broader models that incorporate both prior academic indicators and current student engagement. The present study responds to this gap by examining undergraduate students at Nnamdi Azikiwe University and framing academic risk as a classification problem. Academic performance constitutes the central outcome, while variables such as previous CGPA, admission score, attendance, assessment performance, learning engagement, study hours, age and course load provide potential predictive information. The study therefore shifts attention from explaining academic performance after the outcome has occurred to identifying students who may require support before poor performance becomes entrenched.

The purpose of the study is consequently to evaluate the capacity of selected machine-learning techniques to predict undergraduate academic risk and to determine which student characteristics contribute most strongly to that prediction at Nnamdi Azikiwe University. By comparing alternative classification models and assessing their predictive accuracy, recall, precision, F1-score and discriminatory performance, the study seeks to provide evidence that can support the development of more timely and data-informed academic support systems. Its central problem is that universities may possess substantial student data without having a sufficiently validated, institution-specific predictive framework for identifying academically at-risk undergraduates early enough for meaningful intervention.

Conceptual Literature Review

Machine Learning

Machine learning is generally concerned with developing computer systems that improve their performance by learning from data rather than depending entirely on explicitly programmed rules. Mitchell (1997) describes machine learning as the study of computer algorithms that improve automatically through experience. Similarly, Jordan and Mitchell (2015) explain machine learning as a field concerned with building computers that improve automatically through experience and data.

In this study, machine learning refers to computational methods used to learn patterns from undergraduate students' academic, demographic, behavioural and other relevant records in order to predict their likely academic performance. The concept was measured through the predictive performance of selected algorithms. Accuracy, precision, recall, F1-score and the area under the receiver operating characteristic curve (ROC-AUC) was used to assess the models. These measures will show how well the algorithms distinguish students who are likely to perform satisfactorily from those who may be at risk of poor performance.

Undergraduate Academic Performance

Academic performance refers to the outcomes students achieve from their learning activities. Oxford Bibliographies describes academic achievement as performance outcomes showing the extent to which learners have achieved goals associated with instruction in school, college or university. Akçapınar, Altun and Aşkar (2019) similarly explain academic performance as scores obtained from evaluations of learning activities, which may range from individual learning tasks to semester results and cumulative GPA.

For this study, undergraduate academic performance refers to the level of academic achievement attained by students during their university education. It was measured using students' academic records, particularly semester grades, cumulative grade point average (CGPA), course grades or a derived performance category. Where classification is required, students may be grouped into categories such as satisfactory and at-risk performance using an objectively defined academic threshold. This provides the target variable that the machine learning models will predict.

Early Identification

Early identification in higher education refers to detecting signs of possible academic difficulty before the final academic outcome occurs. Akçapınar et al. (2019) describe educational early-warning systems as an extension of academic performance prediction in which information available

at an earlier stage is used to identify students who may be at risk. Gray and Perkins (2019) also show that early engagement information can be used to identify students who may fail while a course is still in progress.

This study defines early identification as the use of information available during the early stages of an undergraduate student's academic journey to predict subsequent academic performance. The concept was measured by the point at which reliable predictions can be produced, such as using first-semester grades, admission information, early assessment scores, attendance, learning engagement and other available indicators. The aim is not simply to predict performance but to make the prediction early enough to support academic intervention.

At-Risk Students

An at-risk student is a learner whose available academic or behavioural information indicates an increased possibility of poor academic progress, failure or non-completion. Foster and Siddle (2020), for example, examined students identified through learning analytics alerts and assessed whether those alerts could identify students with poorer academic outcomes. Akçapınar et al. (2019) similarly associate early-warning systems with the identification of students who may experience academic failure.

In the present study, an at-risk undergraduate student refers to a student whose observed characteristics and academic records indicate a relatively high likelihood of poor academic performance. The classification will depend on the study's defined performance threshold. Relevant indicators may include previous grades, CGPA, attendance, assessment performance, study behaviour, engagement with learning resources and selected demographic or socioeconomic characteristics. The dependent classification may therefore take a binary form, such as 0 for satisfactory performance and 1 for at-risk performance.

Educational Data Mining

Educational data mining refers to the application of computational and data-mining techniques to information generated within educational environments. Romero and Ventura (2010) describe it as an interdisciplinary area concerned with developing methods for exploring educational data and answering educational questions. Romero and Ventura (2020) further explain that educational data mining and learning analytics have increasingly been applied to educational records to discover patterns that can support learning and educational decision-making.

For this study, educational data mining is understood as the process of extracting useful patterns from undergraduate student data for academic performance prediction. It provides the broader analytical setting within which machine learning is applied. The concept was reflected through the preparation, transformation, selection and analysis of student records before the prediction models are developed. Relevant data may include academic history, admission information, attendance, assessment results, learning engagement and other variables available in the university's records.

Empirical Literature Review

The use of machine learning in higher education has expanded because universities generate substantial amounts of academic and behavioural data. Earlier empirical work established that such information can be used to detect patterns associated with student success or failure. Asif et al. (2017) examined undergraduate student records and used educational data mining to predict academic achievement and study students' academic progression. Their findings demonstrated the usefulness of combining progression information with prediction for understanding student outcomes.

Further evidence was provided by Helal et al. (2018), who considered differences among students when predicting academic performance. Using enrolment information and learning management system activity from an Australian university, the study developed classification models for student outcomes. The findings reinforced the importance of considering student heterogeneity when developing prediction systems.

Similarly, Miguéis et al. (2018) developed an early prediction and segmentation approach using data from 2,459 university students. The study reported predictive accuracy above 95% and found random forest to be stronger than several alternative classification techniques considered. The results indicate that early classification can support strategies aimed at reducing academic failure.

Lau, Sun and Yang (2019) applied an artificial neural network to undergraduate academic performance using several student characteristics and entry-related variables. Their model achieved an overall prediction accuracy of 84.8%. The study demonstrated that neural-network methods can identify complex patterns that conventional statistical approaches may not capture effectively.

Abu Zohair (2019) addressed another practical issue by examining machine learning prediction with a relatively small student dataset. The study found that support vector machine and linear

discriminant analysis produced acceptable classification results. This is relevant where universities have limited historical records and cannot rely on very large datasets.

Also, Gray and Perkins (2019) examined early student engagement as a predictor of academic outcomes. Their results showed that machine learning could identify potentially failing students as early as the fourth week of a semester, with high prediction performance in their institutional setting. The finding supports the value of early rather than end-of-course prediction.

Akçapınar, Altun and Aşkar (2019) used online learning interaction data to investigate whether end-of-term performance could be predicted during earlier weeks. Their study showed that learning traces could support early-warning systems and provide useful information for identifying students who might experience academic difficulty.

Musso, Rodríguez Hernández and Cascallar (2020) extended the application of machine learning beyond simple course grades by predicting GPA, retention and degree completion. Their study showed that predictive models can provide information about several stages of the academic trajectory. This supports the use of machine learning as a broader student-support tool rather than only a grading mechanism.

The importance of early intervention was further examined by Russell, Smith and Larsen (2020). Their study found that students identified as being at risk could use learning analytics information without increasing withdrawal, while engagement with the analytics platform was associated with achieving a passing grade. The finding suggests that prediction can be more useful when linked to appropriate support.

Foster and Siddle (2020) examined learning analytics alerts among first-year undergraduates. They found that no-engagement alerts were more effective than demographic information for identifying students who were not progressing or attaining expected outcomes. This suggests that behavioural learning data may provide stronger early-warning signals than demographic variables alone.

In a Nigerian university context, the application of machine learning by Ekubo and Esiefarienrhe (2022) focused specifically on predicting low academic performance among tertiary students. Using institutional student records and several machine learning techniques, the study demonstrated the feasibility of using educational data mining to classify students according to academic risk.

Yağcı (2022) compared several machine learning algorithms for predicting undergraduate examination performance using midterm grades and programme-related information. The models

achieved classification accuracy of about 70–75%. The study confirms that relatively simple academic information can provide useful predictive signals, although prediction accuracy depends on the variables and learning context used.

Alwarthan (2022) reviewed research on student academic performance prediction and compared machine learning approaches including artificial neural networks, decision trees, support vector machines and Naïve Bayes. The review showed that algorithm performance varies according to the dataset and predictors, indicating that model comparison remains important when developing educational prediction systems.

Christou et al. (2023) similarly developed machine learning models for predicting both academic performance and early student dropout. Their findings support the argument that identifying low-performing students before serious academic difficulty develops can give institutions an opportunity to provide timely support. The study also demonstrates the usefulness of feature engineering in improving predictive models.

Theoretical Framework

Student Involvement Theory

The Student Involvement Theory was propounded by Alexander Astin in 1984. The theory states that students' learning and development are influenced by the amount and quality of physical and psychological energy they invest in their educational experience. Astin argues that students who devote more effort to studying, interacting with lecturers and participating in academic activities are more likely to experience positive educational outcomes.

The theory assumes that student involvement can vary in both quantity and quality, that involvement is closely related to learning, and that educational environments can influence the extent to which students become involved. It also assumes that students' academic outcomes are partly connected to the effort they invest in their educational experience. These assumptions make variables such as attendance, study time, learning platform activity and interaction with academic resources relevant to prediction.

A limitation of the theory is that student involvement is not always easy to measure objectively. Students may also have different opportunities to participate because of financial, social, institutional or personal circumstances. The theory may therefore overlook some structural factors that influence academic outcomes. Despite this limitation, it is relevant to the present study because machine learning can use observable indicators of involvement to identify patterns associated with

academic performance. If students with low participation, weak engagement or limited study activity consistently show higher academic risk, such variables can contribute to early prediction.

Tinto's Student Integration Theory

Tinto's Student Integration Theory was developed by Vincent Tinto in 1975 to explain student persistence and withdrawal from higher education. The theory proposes that students' academic and social integration within the institution affects their commitment to continue their studies. Academic performance, interaction with the academic environment and social relationships therefore form important parts of the student's educational experience.

The theory assumes that students enter higher education with different family backgrounds, previous educational experiences, abilities and commitments. These characteristics interact with the institutional environment. Strong academic and social integration is expected to support persistence, whereas weak integration may increase the possibility of withdrawal or academic difficulty. The theory therefore provides a basis for considering previous academic results, engagement, institutional participation and related student characteristics when predicting outcomes. Nevertheless, the theory is useful to this study because it explains why academic and institutional experiences can produce observable patterns in student records. These patterns can then be learned by predictive algorithms to identify students whose academic outcomes may require attention.

Self-Regulated Learning Theory

Self-Regulated Learning Theory is strongly associated with Barry Zimmerman, whose social cognitive model of self-regulated academic learning was developed from the late 1980s onward. Zimmerman explains self-regulated learning in terms of students' ability to plan, monitor and regulate their learning behaviour, motivation and performance. The theory places particular importance on self-observation, self-judgement, self-reaction and academic self-efficacy.

The theory assumes that learners are active participants in their learning rather than passive recipients of information. It also assumes that students can regulate their cognition, motivation, behaviour and use of learning resources. Effective regulation involves setting goals, monitoring progress, managing study time, selecting appropriate learning strategies and responding to academic feedback. These processes can influence academic achievement.

3.0 METHOD

The study adopts a quantitative predictive research design to examine the possibility of using machine learning techniques to predict undergraduate academic performance and identify students who may be at risk of poor academic outcomes at Nnamdi Azikiwe University. The design is appropriate because the study relies on existing student-level data and focuses on identifying patterns that can be used to make predictions about subsequent academic performance.

Unlike a conventional explanatory study that mainly estimates the effect of individual variables on an outcome, this study focuses on prediction. Several student characteristics are therefore considered jointly rather than examining each variable in isolation. The approach allows different machine learning algorithms to be trained, tested and compared based on their ability to correctly classify students according to their academic performance.

Study Area and Population

The study is conducted at Nnamdi Azikiwe University, Awka, Anambra State, Nigeria. The university provides an appropriate setting for the study because it has a large undergraduate population and generates academic records that can be used for predictive analysis. The population comprises undergraduate students whose relevant academic and student information is available for the period covered by the study.

The unit of analysis is the individual undergraduate student. The study uses student-level observations rather than aggregate departmental or faculty-level information. This makes it possible to identify differences in academic performance among individual students and develop a model capable of identifying students who may be at risk.

Data Source and Data Collection

The study relies primarily on secondary student-level data obtained from relevant academic and administrative records of Nnamdi Azikiwe University. Depending on availability, the dataset contain information on students' previous academic performance, admission characteristics, course or programme information, attendance, assessment results, learning engagement and other relevant characteristics.

The use of existing institutional records reduces reliance on self-reported information and allows the prediction models to learn from actual student outcomes. Only variables that are available before the target academic outcome observed was used as predictors. This is important because using

information that becomes available after the outcome would create data leakage and produce misleadingly high prediction accuracy.

Definition and Measurement of Variables

The dependent variable is undergraduate academic performance. It represents the academic outcome that the machine learning models are expected to predict. Academic performance may be measured using students' CGPA, semester GPA, final course performance or another academically appropriate institutional measure.

For the early-identification component, students were classified into two groups based on a predetermined academic performance threshold. Students whose performance falls below the specified threshold were classified as at-risk, while students meeting or exceeding the threshold were classified as not at-risk. The threshold was determined from the university's grading system and the objective of the study rather than being selected after examining model results.

The explanatory variables consist of student characteristics that are theoretically and empirically related to academic performance. These include previous GPA or CGPA, admission score, attendance, continuous assessment performance, examination scores available at the prediction stage, course load, programme of study, level of study, gender, age and learning engagement indicators. The final set of predictors will depend on the completeness and quality of the available institutional data.

Data Pre-processing

Data preprocessing were undertaken before model development. The process involves checking the dataset for missing values, duplicate observations, inconsistent entries and extreme values. Variables were coded into formats suitable for machine learning algorithms, while categorical variables were appropriately encoded.

Missing observations were treated using an appropriate method depending on their extent and pattern. Variables with excessive missing information may be excluded where their inclusion could reduce the reliability of the analysis. Numerical variables may also be standardized where required by the selected algorithms.

Particular attention were given to class imbalance. If the number of at-risk students is substantially lower than the number of students classified as not at-risk, the models may become biased towards the majority class. Appropriate techniques such as class weighting or resampling may

therefore be considered. These procedures were applied only to the training data to prevent information from the test set entering the model-development process.

Machine Learning Models

The study compared selected supervised machine learning algorithms for predicting undergraduate academic performance. The choice of multiple algorithms is important because different algorithms identify patterns in different ways and no single technique is expected to perform equally well across all datasets.

The candidate models include logistic regression, decision tree, random forest, support vector machine and an artificial neural network. Logistic regression provided a useful baseline model because it is relatively simple and its coefficients can be interpreted. Decision tree models can capture non-linear relationships and interactions among predictors. Random forest combines several decision trees and can improve prediction by reducing the dependence on a single tree.

Support vector machine was considered because of its ability to identify complex boundaries between classes, while the artificial neural network was used to examine whether a more flexible non-linear learning structure provides better predictive performance. The final models included in the analysis depended on data quality, sample size and computational suitability.

Model Development

The dataset was divided into training and testing portions. The training data was used to develop the predictive models, while the test data was kept separate for evaluating their performance on previously unseen observations. Where the sample size permitted, cross-validation was employed within the training data to improve the reliability of model selection and hyperparameter tuning.

The general classification problem can be expressed as:

$Yi = f(X_1i, X_2i, X_3i,...,Xki) + \varepsilon i$

where Yi represents the academic risk status of student i, X1,…,Xk represent the selected student characteristics and academic indicators, and f(.) represents the machine learning function used to learn the relationship between the predictors and academic outcome.

For the binary classification model, the outcome is represented as:

Yi={1,if student i is classified as at-risk, 0,if student i is classified as not at- }

The models learned the relationship between the available predictors and this outcome using the training dataset. The resulting models was then applied to the test observations without using their known outcomes during prediction.

Model Training and Hyperparameter Tuning

Each selected algorithm was trained using the same training data and evaluated under comparable conditions. Where appropriate, model parameters was tuned using cross-validation. This process helped to determine parameter combinations that improve predictive performance without simply fitting the training observations.

The tuning process was restricted to the training data. The test dataset remained untouched until the final evaluation. This separation is important because using the test data during model development can lead to overly optimistic estimates of predictive accuracy.

Evaluation of Predictive Performance

The performance of the machine learning models were evaluated using a confusion matrix and several complementary performance measures. Accuracy indicated that the proportion of all the observations were correctly classified. However, because the study focused on identifying at-risk students, accuracy alone may be inadequate, particularly where the two classes are imbalanced.

Precision measured the proportion of students classified as at-risk who are actually at-risk, while recall measured the proportion of actual at-risk students correctly identified by the model. The F1-score provided a combined measure of precision and recall.

The measures is expressed as follows:

Accuracy = TP + TNTP + TN + FP + FNAccuracy = \frac {TP+TN}{TP+TN+FP+FN}

Precision = TP + FPPrecision =\frac{TP}{TP+FP}

Recall =TP + FNRecall = \frac{TP}{TP+FN} F1 = 2 (Precision × Recall) Precision + Recall F1 =\frac{2(Precision\times Recall)}{Precision + Recall}

where TP represents true positives, TN true negatives, FP false positives and FN false negatives.

The receiver operating characteristic curve and the area under the curve (ROC-AUC) was also considered. A higher ROC-AUC indicates stronger ability of the model to distinguish between students classified as at-risk and those classified as not at-risk. Since the purpose of the study is early identification, particular attention was given to recall and F1-score alongside overall accuracy.

Comparative Model Assessment

The predictive performance of the competing algorithms were compared using the same evaluation criteria and test observations. The comparison focused on whether the models can reliably distinguish students who are likely to experience poor academic outcomes from those who are not.

Rather than relying on accuracy alone, the study considered the combined evidence from precision, recall, F1-score, ROC-AUC and the confusion matrix. This is particularly important because failing to identify an actual at-risk student may have greater practical consequences for an early-warning system than incorrectly flagging a student who subsequently performs well.

Ethical Considerations

Because the study involves student-level academic information, confidentiality and responsible use of data are important. Student names, identification numbers and other direct identifiers were not be included in the analytical dataset. The student records were anonymised before analysis.

The data was used strictly for academic research purposes. Access to institutional records were obtained through the appropriate university authorities, and the study complied with relevant institutional requirements concerning student privacy and research ethics. The prediction results were interpreted as indicators of academic risk rather than permanent labels of students' abilities or future potential.

Software and Analytical Procedure

Data preparation and analysis was conducted using appropriate statistical and machine learning software such as Python, R, SPSS and other suitable analytical packages. Python-based machine learning libraries such as scikit-learn was used for preprocessing, model development, cross-validation and performance evaluation.

The analysis follows a sequential process beginning with data cleaning and exploratory analysis, followed by variable preparation and feature selection. The dataset are then be divided into training and testing samples. The selected algorithms was trained and tuned using the training data, after which their predictive performance was assessed using the independent test data.

The final stage involved comparing the models and identifying the model or models that provide reliable early prediction of academic risk. The results was presented using appropriate tables, performance measures and, where useful, confusion matrices and ROC curves.

RESULTS AND DISCUSSION

Sample profile

The dataset held 400 students, and 99 of them (24.8%) fell in the at-risk group. That left 301 in good standing. Table 1 puts the two groups side by side.

Table 1.

Characteristics of students by academic standing

Variable	Statistic	Overall	At risk	Not at risk	p
Students, n (%)		400 (100)	99 (24.8)	301 (75.2)	
UTME score	Mean (SD)	233.07 (20.40)	220.25 (19.47)	237.29 (18.90)	< .001
Post-UTME score	Mean (SD)	57.52 (11.03)	51.55 (10.29)	59.52 (10.55)	< .001
O'level points	Mean (SD)	26.79 (4.23)	24.40 (4.09)	27.57 (3.97)	< .001
Continuous assessment (%)	Mean (SD)	57.95 (12.88)	49.15 (10.69)	60.90 (12.19)	< .001
Attendance (%)	Mean (SD)	70.86 (14.58)	65.30 (13.91)	72.67 (14.36)	< .001
First-semester GPA	Mean (SD)	3.04 (0.62)	2.43 (0.51)	3.24 (0.52)	< .001
Carryover courses	Mean (SD)	1.41 (1.23)	2.51 (1.13)	1.05 (1.03)	< .001
Age (years)	Mean (SD)	20.44 (1.33)	20.76 (1.49)	20.33 (1.26)	.012
Gender	Female	233 (58.2)	55 (55.6)	178 (59.1)	.611
	Male	167 (41.8)	44 (44.4)	123 (40.9)	
Mode of entry	Direct Entry	52 (13.0)	13 (13.1)	39 (13.0)	1.000
	UTME	348 (87.0)	86 (86.9)	262 (87.0)	

Note. p-values come from Welch t-tests (numeric variables) and chi-square tests (categorical variables). Missing values were filled after the split, so this table uses the observed values only.

Author’s computation (2026)

Table 1 shows that almost every academic variable separated the groups clearly ($p < .001$). At-risk students came in with lower UTME scores ($M = 220.25$ against 237.29) and lower post-UTME scores (51.55 against 59.52), and their O'level points were lower as well. Once in the system the gap

grew. Their continuous assessment average sat about 12 points below that of their peers (49.15 against 60.90), they attended less (65.30% against 72.67%) and they carried more failed courses, 2.51 on average against 1.05. First-semester GPA showed the widest contrast, 2.43 against 3.24. Age differed only slightly (20.76 against 20.33, p = .012). Gender (p = .611) and mode of entry (p = 1.000) did not differ at all, so risk was not tied to either in this sample.

Only about one student in four was at risk, so SMOTE was applied inside the training folds. The test set was left alone. Training had 320 students (79 at risk) and testing 80 students (20 at risk).

4.2 Tuning and cross-validated performance

Table 2 lists the settings picked by grid search. Table 3 gives the cross-validated scores on the training set.

Table 2

Hyperparameters selected by grid search

Model	Selected hyperparameters	Mean CV ROC-AUC
Logistic Regression	C = 0.1	0.911
Random Forest	max_depth = 5; min_samples_leaf = 1; n_estimators = 500	0.899
SVM (RBF)	C = 0.1; gamma = 0.01	0.91
XGBoost	learning_rate = 0.05; max_depth = 3; n_estimators = 200; subsample = 0.8	0.887

Author's computation (2026)

Table 3

Ten-fold cross-validated performance on the training set, mean (SD)

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC
Logistic Regression	0.828 (0.067)	0.631 (0.140)	0.784 (0.160)	0.692 (0.128)	0.911 (0.063)
Random Forest	0.859 (0.058)	0.700 (0.130)	0.786 (0.147)	0.733 (0.116)	0.899 (0.067)
SVM (RBF)	0.834 (0.075)	0.635 (0.149)	0.848 (0.156)	0.719 (0.131)	0.910 (0.063)
XGBoost	0.834 (0.067)	0.658 (0.136)	0.696 (0.149)	0.674 (0.135)	0.887 (0.062)

Author's computation (2026)

Table 3 shows that the four models finished close together. ROC-AUC went from 0.887 for XGBoost to 0.911 for logistic regression, which is good discrimination by the usual rule of thumb. Random forest led on accuracy (0.859), precision (0.700) and F1 (0.733), and that is why it was chosen as the best model, before the test set was opened. The SVM caught the most at-risk students in the folds (recall 0.848) but raised more false alarms to do it, with precision at 0.635. XGBoost was weakest on recall, 0.696.

The spread across folds was not small. Standard deviations for F1 ran from 0.116 to 0.135. That is what is expected when each fold holds 32 students and only around eight of them are at risk, so the gaps in Table 3 should be read loosely.

4.3 Performance on the test set

Table 4 shows how the tuned models did on the 80 students they had never seen.

Table 4

Test-set performance of the models

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC
Logistic Regression	0.812	0.593	0.8	0.681	0.903
Random Forest	0.762	0.52	0.65	0.578	0.889
SVM (RBF)	0.825	0.615	0.8	0.696	0.911
XGBoost	0.775	0.536	0.75	0.625	0.897

Author's computation (2026)

In Table 4, the order changed here. Random forest, the choice from cross-validation, had the lowest test scores: accuracy 0.762, recall 0.650 and F1-score of 0.578. The SVM did best on accuracy (0.825), F1 (0.696) and AUC (0.911), with logistic regression close behind (F1 0.681, AUC 0.903). With only 20 at-risk students in the test set, one student moves recall by five points, and a swing this size can come from the luck of the split. The steadier finding is that every test AUC stayed between 0.889 and 0.911, in line with the cross-validated values. Nothing fell apart on new data.

Table 5

Confusion matrix counts on the test set

Model	At risk caught (TP)	At risk missed (FN)	Wrongly flagged (FP)	Correctly cleared (TN)
Logistic Regression	16	4	11	49

Model	At risk caught (TP)	At risk missed (FN)	Wrongly flagged (FP)	Correctly cleared (TN)
Random Forest	13	7	12	48
SVM (RBF)	16	4	10	50
XGBoost	15	5	13	47

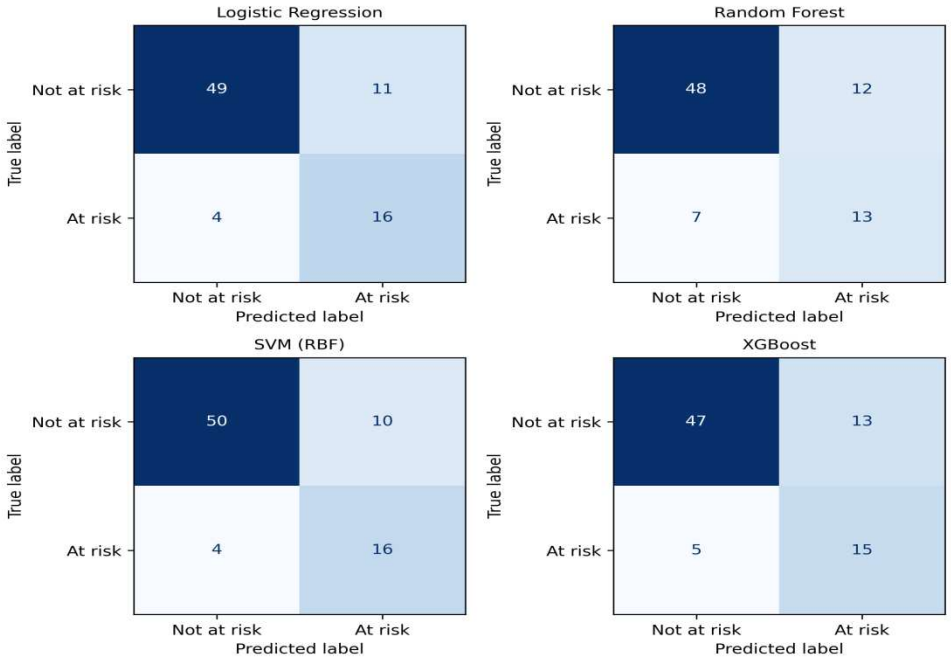

Figure 1. Confusion matrices for the four models on the test set

Author's computation (2026)

Table 5 and Figure 1 show the kind of mistakes each model made. Logistic regression and the SVM each caught 16 of the 20 at-risk students, while random forest caught only 13 and missed 7. False alarms ran from 10 (SVM) to 13 (XGBoost). For a tool meant to trigger early help, a missed student is the costlier error, because that student gets no support. A student flagged by mistake usually just ends up in an extra advising meeting. By that reasoning the SVM and logistic regression seem the better choices in this run.

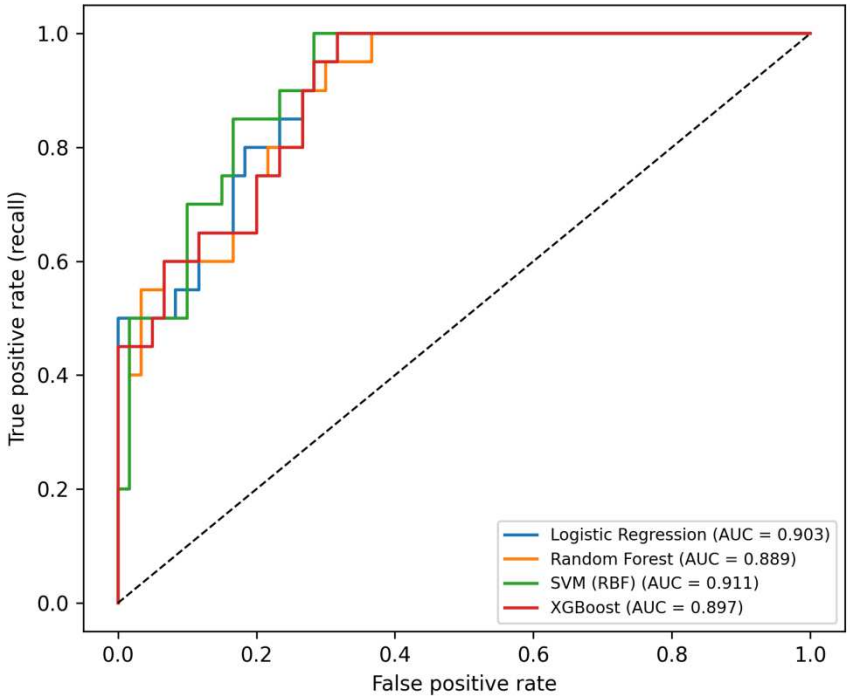

Figure 2. ROC curves of the four models on the test set

Author's computation (2026)

The ROC curves in Figure 2 bunch together well above the diagonal. It is hard to pull them apart by eye, which fits the small differences in AUC.

Are the differences between models real?

Fold-wise F1 scores were compared to see whether the gaps in Table 3 could be chance. The Friedman test across the four models was significant, $\chi^2(3) = 10.08$, $p = .018$, so at least one model differed from the rest. Table 6 compares the best model, random forest, with each of the others.

Table 6.

Pairwise comparison of random forest with the other models (fold-wise F1)

Comparison	Mean difference	Test	Statistic	p-value	Significant at 5%
Random Forest vs Logistic Regression	0.041	Wilcoxon W	1.0	0.031	Yes
Random Forest vs SVM (RBF)	0.014	Wilcoxon W	16.5	0.516	No
Random Forest vs XGBoost	0.059	Wilcoxon W	0.0	0.004	Yes

Author's computation (2026)

Random forest beat logistic regression by 0.041 (p = .031) and XGBoost by 0.059 (p = .004), but could not be told apart from the SVM (difference 0.014, p = .516). Given what happened on the test set, though, I would not crown it. Scores from overlapping folds are not independent, and tests like these tend to be a little too generous (Nadeau & Bengio, 2003). They are best treated as supporting evidence, and a bigger sample would settle the matter better than another test.

Which variables matter most: random forest importance

Table 7

Random forest feature importance (top 12 predictors)

Rank	Feature	Importance
1	First-semester GPA	0.317
2	Carryover courses	0.2261
3	Continuous assessment (%)	0.1094
4	UTME score	0.0815
5	Post-UTME score	0.0755
6	O'level points	0.0533
7	Attendance (%)	0.0451
8	Age (years)	0.0341
9	Gender (female)	0.0175
10	Gender (male)	0.0157
11	Entry mode (direct entry)	0.0127
12	Entry mode (UTME)	0.0121

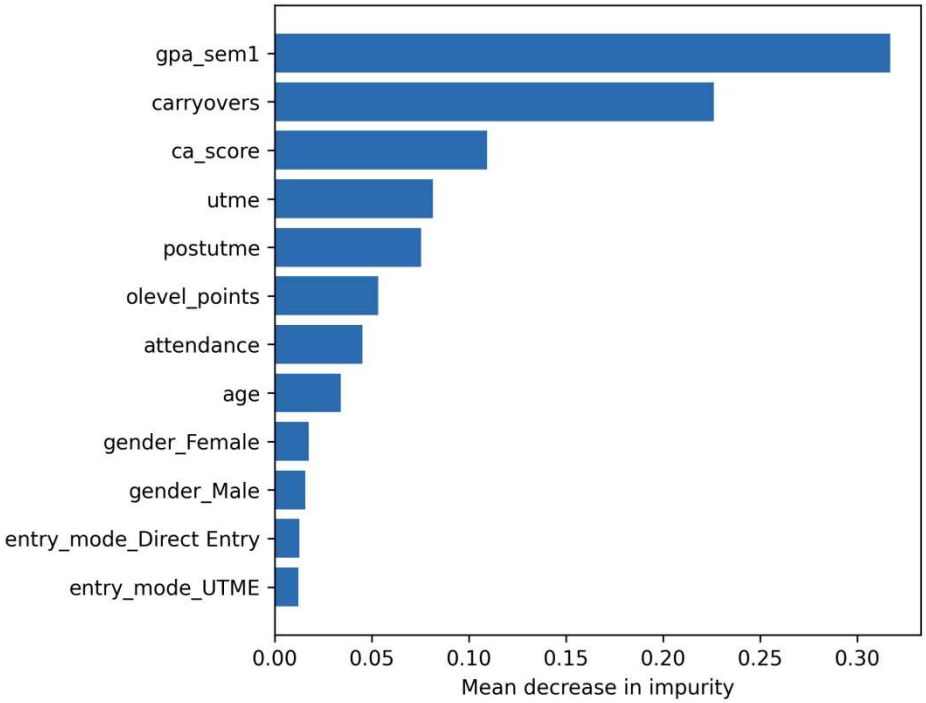

Figure 3. Random forest feature importance for the predictors

Author's computation (2026)

Figure 3 shows that first-semester GPA came out on top (0.317), with carryover courses second (0.226). Those two together carry about 54% of the total importance, and adding continuous assessment (0.109) takes the top three to roughly 65%. The entry scores are further down: UTME 0.082, post-UTME 0.076 and O'level points 0.053, about 21% between them. Attendance (0.045) and age (0.034) add a little. Gender and mode of entry are each under 0.02.

So what students do in their first semester says more about their risk than what they walked in with. Attendance is the odd one. It differed clearly between the groups in Table 1, yet ranks low here, probably because much of what it tells the model is already in GPA and continuous assessment. Impurity-based importance also leans toward variables with many distinct values and does not show direction, which is why SHAP was run as well.

Reading the model with SHAP

Table 8

Mean absolute SHAP values for the random forest model (top 12 predictors)

Rank	Feature	Mean |SHAP|

Rank	Feature	Mean |SHAP|
1	First-semester GPA	0.1427
2	Carryover courses	0.1221
3	Continuous assessment (%)	0.0552
4	UTME score	0.0444
5	Post-UTME score	0.0363
6	O'level points	0.0306
7	Attendance (%)	0.0251
8	Age (years)	0.0185
9	Entry mode (direct entry)	0.0078
10	Gender (female)	0.006
11	Entry mode (UTME)	0.006
12	Gender (male)	0.0051

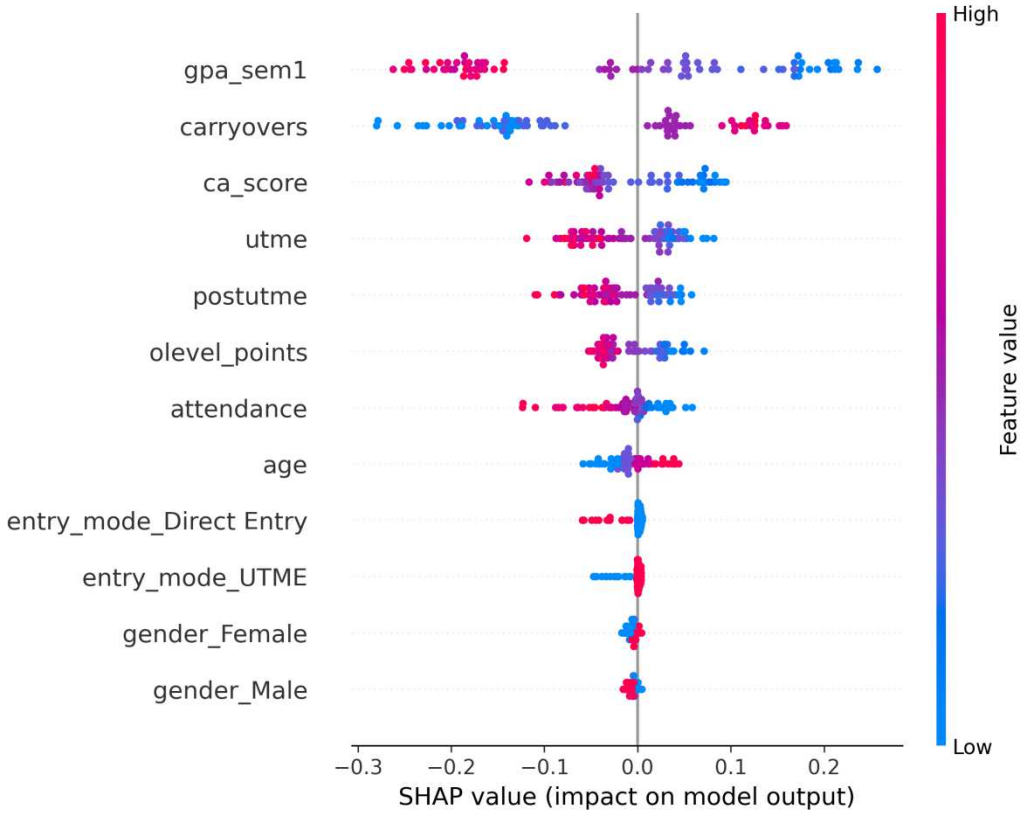

Figure 4. SHAP summary plot for the random forest model. Positive values push a student toward the at-risk class.

Author's computation (2026)

The SHAP ranking in Table 8 is nearly the same as the importance ranking: first-semester GPA (0.143), carryovers (0.122) and continuous assessment (0.055) lead, followed by UTME (0.044), post-UTME (0.036), O'level points (0.031), attendance (0.025) and age (0.019).

Figure 4 adds the direction. Each dot is one student, red for a high value of the variable and blue for a low one. Low GPA sits on the right, pushing toward at risk, and high GPA pulls left. Carryovers work the other way round: many failed courses push a student strongly to the right. Low continuous assessment, low entry scores and low attendance all push right too. Age has a small effect, with older students nudged slightly toward risk, and direct-entry students nudged slightly away from it. Gender hardly moves the prediction at all, which agrees with the null result in Table 1. Nothing here looks strange. The model leans on the things a lecturer would name if asked who is struggling, and that makes it easier to trust.

Discussion of Results and Findings

The findings show that academic risk was strongly associated with students' academic history and their performance during the early part of university education. Of the 400 students studied, 99 (24.8%) were classified as at risk, while 301 (75.2%) were not. This means that about one in every four students in the sample was potentially vulnerable to poor academic outcomes. The result points to the value of identifying academic difficulty before it becomes more difficult to address.

Clear differences were observed between the two groups. At-risk students recorded lower UTME scores, post-UTME scores, O'level points, continuous assessment scores, attendance and first-semester GPA. They also had more carryover courses. All these academic differences were statistically significant at the 1% level. First-semester GPA showed a particularly wide gap, with an average of 2.43 among at-risk students compared with 3.24 among those not at risk. Carryover courses also differed considerably, averaging 2.51 for at-risk students and 1.05 for their counterparts. In contrast, gender and mode of entry were not significantly different between the two groups. This suggests that academic progress and engagement provided more useful information about risk than these background characteristics in the present sample.

These results agree with previous evidence that earlier academic performance is an important source of predictive information. Yağcı (2022) found that previous examination performance could be used to predict undergraduate achievement, while Çırak, Akıllı and Ekinci (2024) identified previous academic performance among the important predictors of student outcomes. The present findings also support Foster and Siddle (2020), who showed that engagement-related information

could help identify students with weaker academic outcomes. The lower attendance and continuous assessment scores observed among at-risk students suggest that signs of academic difficulty may be visible before final examination results are available.

The machine-learning results further support the usefulness of predictive modelling. During ten-fold cross-validation, all four models produced ROC-AUC values above 0.88. Logistic regression recorded the highest cross-validated AUC (0.911), followed by SVM (0.910), random forest (0.899) and XGBoost (0.887). Random forest recorded the highest cross-validated accuracy, precision and F1-score. This is broadly consistent with earlier studies showing that machine learning can identify students who may experience academic difficulty. Miguéis et al. (2018), Gray and Perkins (2019), and Akçapınar, Altun and Aşkar (2019) all reported useful predictive performance using machine-learning approaches in different higher-education settings.

An important finding, however, was the difference between cross-validation and independent test performance. Random forest appeared strongest during cross-validation but produced the weakest test performance among the four models, with accuracy of 0.762, recall of 0.650 and F1-score of 0.578. On the test data, SVM produced the highest accuracy (0.825), F1-score (0.696) and ROC-AUC (0.911), while logistic regression also performed strongly with an accuracy of 0.812, recall of 0.800 and AUC of 0.903. The difference suggests that internal validation results should not be treated as proof that one algorithm will always perform better on new observations.

The confusion matrices make this point clearer. SVM and logistic regression each correctly identified 16 of the 20 at-risk students, compared with 13 for random forest and 15 for XGBoost. SVM generated 10 false positives, while logistic regression generated 11. For an early-warning system, detecting students who are actually at risk is important because a missed student may not receive timely support. At the same time, the test sample contained only 20 at-risk students, meaning that the classification of a few students could noticeably change recall and F1-score. The findings should therefore be interpreted with some caution.

The feature-importance and SHAP results provide further insight into what drove the predictions. First-semester GPA was the strongest predictor, followed by carryover courses and continuous assessment. These three variables accounted for a large share of the model's predictive information. UTME, post-UTME and O'level scores contributed less, while gender and mode of entry had very small contributions. The SHAP analysis also showed that lower GPA, more carryover courses, lower continuous assessment scores, lower admission scores and lower attendance generally pushed predictions towards the at-risk category.

This pattern agrees with Astin's (1984) Student Involvement Theory and the Self-Regulated Learning Theory associated with Zimmerman. Lower attendance and weaker continuous assessment performance can reflect differences in students' involvement, study behaviour and regulation of learning. The findings also have some connection with Tinto's Student Integration Theory (1975), since continuing academic difficulty may reflect weaker academic integration. However, the present study did not directly measure social integration, so the theory provides broader context rather than a direct test of all its propositions.

The results differ somewhat from studies reporting prediction accuracies above 90%, such as Miguéis et al. (2018) and Gray and Perkins (2019). Such differences are understandable because the studies used different institutions, samples, variables, prediction periods and definitions of academic risk. Some relied heavily on learning-management-system activity, whereas the present study combined admission information with academic records and selected engagement indicators. The smaller sample of 400 students may also have contributed to greater variation in model performance. The use of SMOTE to address the 24.8% at-risk class is another factor that may influence model behaviour, although it was applied only within the training folds.

The findings have practical implications for Nnamdi Azikiwe University. An early-warning system could give greater attention to first-semester GPA, carryover courses and continuous assessment rather than relying mainly on admission characteristics. Attendance can also be monitored as an additional signal. Students identified as potentially at risk could then be referred for academic advising, tutorials, mentoring or other forms of support. Such predictions should be treated as early signals rather than permanent labels, since students' circumstances and performance can change.

The study addresses a specific gap in the literature by providing institution-specific evidence from Nnamdi Azikiwe University. Much of the existing research has focused on institutions outside Nigeria, online learning environments, individual courses or particular student groups. Nigerian evidence remains relatively limited and is concentrated in specific institutional settings. By comparing logistic regression, random forest, SVM and XGBoost using undergraduate student records and independently testing their predictions, this study extends the evidence on machine-learning-based academic risk prediction within the Nigerian university context. The findings show that machine learning can provide useful early-warning information, but they also demonstrate the importance of independent testing, larger datasets and validation across different student cohorts and academic programme

Conclusion

This study set out to see whether machine learning could pick out undergraduates likely to end up in academic trouble, early enough for someone to do something about it. The study concludes that machine learning can provide a useful approach for identifying undergraduate students who are academically at risk at Nnamdi Azikiwe University. The results show clear differences between at-risk and non-at-risk students, particularly in first-semester GPA, carryover courses, continuous assessment, attendance, and admission scores. First-semester GPA and carryover courses emerged as the strongest predictors, while gender and mode of entry contributed very little to prediction. The models also recorded good discriminatory ability, with test-set ROC-AUC values ranging from 0.889 to 0.911. Although Random Forest performed strongly during cross-validation, its performance declined on the test data, while SVM and Logistic Regression performed relatively well on unseen students. This suggests that model selection should not depend on a single performance measure or training result. Overall, the findings support the use of machine learning as an early-warning tool for academic intervention. However, the small test sample and the inclusion of first-semester GPA mean that the findings should be interpreted with some caution and validated with larger student datasets.

Recommendations

Based on the results, the following recommendations can be made:

1. The University should strengthen early academic monitoring by paying closer attention to students' GPA, continuous assessment, attendance and carryover courses, as these characteristics clearly distinguished at-risk students.
2. Nnamdi Azikiwe University should adopt machine learning-based early warning systems to identify students at risk and provide timely academic advising and support.
3. Greater attention should be given to first-semester GPA, carryover courses and continuous assessment because they were the most important predictors of academic risk in the models.
4. The University should evaluate model performance using multiple measures and repeated cross-validation rather than relying on accuracy alone, particularly because the test sample was small and model rankings differed between training and test results.

References

Abu Zohair, L. M. (2019). Prediction of student's performance by modelling small dataset size. *International Journal of Educational Technology in Higher Education, 16*, Article 27. https://doi.org/10.1186/s41239-019-0160-3

Akçapınar, G., Altun, A., & Aşkar, P. (2019). Using learning analytics to develop early-warning system for at-risk students. *International Journal of Educational Technology in Higher Education, 16*, Article 40. https://doi.org/10.1186/s41239-019-0172-z

Alwarthan, S. A., Aslam, N., & Khan, I. U. (2022). Predicting student academic performance at higher education using data mining: A systematic review. *Applied Computational Intelligence and Soft Computing, 2022*, Article 8924028. https://doi.org/10.1155/2022/8924028

Asif, R., Merceron, A., Ali, S. A., & Haider, N. G. (2017). Analyzing undergraduate students' performance using educational data mining. *Computers & Education, 113*, 177–194. https://doi.org/10.1016/j.compedu.2017.05.007

Astin, A. W. (1984). Student involvement: A developmental theory for higher education. *Journal of College Student Personnel, 25*(4), 297–308.

Christou, V., Tsoulos, I., Loupas, V., Tzallas, A. T., Gogos, C., Karvelis, P. S., Antoniadis, N., Glavas, E., & Giannakeas, N. (2023). Performance and early drop prediction for higher education students using machine learning. *Expert Systems with Applications, 225*, 120079. https://doi.org/10.1016/j.eswa.2023.120079

Çırak, C. R., Akıllı, H., & Ekinci, Y. (2024). Development of an early warning system for higher education institutions by predicting first-year student academic performance. *Higher Education Quarterly, 78*(4), Article e12539. https://doi.org/10.1111/hequ.12539

Ekubo, E. A., & Esiefarienrhe, B. M. (2022). Using machine learning to predict low academic performance at a Nigerian university. *The African Journal of Information and Communication, 30*, 1–33. https://doi.org/10.23962/ajic.i30.14839

Foster, E., & Siddle, R. (2020). The effectiveness of learning analytics for identifying at-risk students in higher education. *Assessment & Evaluation in Higher Education, 45*(6), 842–854.

Gray, C. C., & Perkins, D. (2019). Utilizing early engagement and machine learning to predict student outcomes. *Computers & Education, 131*, 22–32. https://doi.org/10.1016/j.compedu.2018.12.006

Helal, S., Li, J., Liu, L., Ebrahimie, E., Dawson, S., Murray, D. J., & Long, Q. (2018). Predicting academic performance by considering student heterogeneity. *Knowledge-Based Systems, 161*, 134–146. https://doi.org/10.1016/j.knosys.2018.07.042

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science, 349*(6245), 255–260. https://doi.org/10.1126/science.aaa8415

Lau, E. T., Sun, L., & Yang, Q. (2019). Modelling, prediction and classification of student academic performance using artificial neural networks. *SN Applied Sciences, 1*, Article 982. https://doi.org/10.1007/s42452-019-0884-7

Miguéis, V. L., Freitas, A., Garcia, P. J. V., & Silva, A. (2018). Early segmentation of students according to their academic performance: A predictive modelling approach. *Decision Support Systems, 115*, 36–51. https://doi.org/10.1016/j.dss.2018.09.001

Mitchell, T. M. (1997). *Machine learning*. McGraw-Hill.

Musso, M. F., Rodríguez Hernández, C. F., & Cascallar, E. C. (2020). Predicting key educational outcomes in academic trajectories: A machine-learning approach. *Higher Education, 80*, 875–894. https://doi.org/10.1007/s10734-020-00520-7

Nadeau, C., & Bengio, Y. (2003). Inference for the generalization error. *Machine Learning, 52*, 239–281. https://doi.org/10.1023/A:1024068626366

Nadeau, C., & Bengio, Y. (2003). Inference for the generalization error. Machine Learning, 52(3), 239–281

Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 40*(6), 601–618.

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

Russell, J.-E., Smith, A., & Larsen, R. (2020). Elements of success: Supporting at-risk student resilience through learning analytics. *Computers & Education, 152*, 103890. https://doi.org/10.1016/j.compedu.2020.103890

Steinmayr, R., Meißner, A., Weidinger, A. F., & Wirthwein, L. (2014). Academic achievement. In *Oxford Bibliographies in Education*. Oxford University Press. https://doi.org/10.1093/obo/9780199756810-0108

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*(1), 89–125. https://doi.org/10.3102/00346543045001089

Yağcı, M. (2022). Educational data mining: Prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments, 9*, Article 11. https://doi.org/10.1186/s40561-022-00192-z

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology, 81*(3), 329–339. https://doi.org/10.1037/0022-0663.81.3.329

UNESCO. (2024). *2024/5 Global Education Monitoring Report: Monitoring SDG 4*. United Nations Educational, Scientific and Cultural Organization.

.