World Journal of SDGs Review
Vol.2 (2025), Pp.1-8
ISSN (Online): 3105-4463, ISSN (Print): 3105-4455
DOI: https://doi.org/10.71016/wjsdgr/47t3ee13

Original Article https://hnpublisher.com

Predictive Modeling of Undergraduate Academic Performance Using XGBoost and Implications for Educational Policy in Nigeria

Moses Adekunle Ogunrinde[1], Eugenia A. Okwilagwe[2]

[1]PhD Candidate, Institute of Education, University of Ibadan, Ibadan, Oyo State, Nigeria.
[2]Professor of Social Science Education, Institute of Education, University of Ibadan, Nigeria.
Correspondence: geniaokwilagwe2004@yahoo.com[2]

ABSTRACT

Aim of the Study: Previous studies have paid little attention on the study that investigates the application of the XGBoost machine learning algorithm in predicting undergraduate students' academic performance, specifically focusing on final degree classification outcomes in Nigerian universities.

Methodology: Drawing on a dataset comprising academic, socio-economic, and institutional variables from 1,200 students, the study developed and validated a predictive model using supervised learning techniques. The model achieved an accuracy of 43.0%, an F1 score of 0.407, and a ROC AUC of 0.642, indicating fair discriminatory power but moderate overall predictive performance. This study therefore employed a quantitative, predictive research design using machine learning techniques to model and forecast undergraduate academic performance outcomes in Nigerian universities.

Findings: According to the reuslts, it was discovered that while XGBoost shows potential for identifying performance trends and at-risk students, its effectiveness is constrained by limitations in data diversity, class imbalance, and contextual complexity.

Conclusion: The study contributes to the emerging field of educational data mining in sub-Saharan Africa and highlights the importance of integrating machine learning tools into academic planning and decision-making. It recommends improvements in data collection practices, the inclusion of behavioural and contextual variables, and the adoption of hybrid modeling approaches to enhance accuracy.

Keywords: Predictive modeling, Undergraduate, Academic performance, Educational policy, Nigeria.

Article History

Received:
July 25, 2025

Revised:
October 01, 2025

Accepted:
October 02, 2025

Online:
October 03, 2025

1. INTRODUCTION

Over the past few years, there has been a growing access to educational data that has presented new avenues to enhancing learning outcomes based on data driven decision-making. Among the most urgent problems of higher education in Nigeria is the necessity to understand and address the complexity of the factors that influence the academic performance of students and the degree classification outcomes. Although these problems have been studied by conventional statistical approaches, increased complexity and the sheer amount of educational data require more advanced analytical tools to reveal the hidden trends and aid in early intervention approaches.

The field of machine learning as a subset of artificial intelligence has turned out to be a potent instrument of educational data mining, enabling researchers and educators to model with greater predictive accuracy the results of academic performance (Nghe, Janecek, and Haddawy, 2007). Extreme Gradient Boosting (XGBoost) is one of the well-known and predominant machine learning algorithms because of its capacity to process non-linear data structures, missing values, and generate powerful predictions in complicated classification scenarios (Chen and Guestrin, 2016). Nevertheless, its introduction in the higher education of Nigeria is comparatively minimal, despite its commonly used application in the systems of global education.

This paper will discuss how the XGBoost algorithm can be used to predict the academic performance of undergraduate students in Nigerian universities, and specifically, the results of the undergraduate degree classification. The model was tested using a verifiably validated data set with several performance measures such as accuracy, F1 score, and ROC AUC. The initial results show that the model has a reasonable discriminatory power (ROC AUC = 0.642), but its total predictive accuracy (0.430) and F1 score (0.407) demonstrate what can be viewed as significant shortcomings in the model attempting to capture the full range of student performance.

Findings are not only adding to the existing body of literature about machine learning in education, but are also putting forth the dire necessity of integrating data science into education planning and policy in Nigeria. Knowledge of these predictive trends may benefit academic advising, curriculum development and early intervention on the at risk students and thus improve the quality and equity of higher education. There are also implications of this to institutional policy, early interventions, and national education reform. The study highlights the importance of predictive analytics in facilitating equity, efficiency, and success of Nigerian higher education institutions.

2. LITERATURE REVIEW

Educational applications of machine learning (ML) methods have become increasingly popular in recent decades, providing methods to predict, classify, and diagnose academic behaviours and outcomes of students. Many studies have shown the possibility of predictive analytics being useful in higher education to aid early intervention and policy change. The success of such models however, greatly depends on the choice of algorithms, quality of data and contextual relevance, especially in developing nations such as Nigeria.

Moreover, machine learning models are increasingly used to forecast students' academic performance based on demographic, academic, and behavioral features. Nghe, Janecek, and Haddawy (2007) conducted one of the early comprehensive comparisons of ML algorithms in educational data mining. They evaluated decision trees, Naïve Bayes, and support vector machines and found that no single model consistently outperformed the others across datasets, although ensemble methods showed promising results. Similarly, Kotsiantis, Pierrakeas, and Pintelas (2004) emphasized that combining models or using ensemble techniques often leads to improved predictive outcomes in distance and conventional learning systems.

Besides, among the ensemble learning techniques, XGBoost has emerged as a particularly effective algorithm due to its scalability, regularization, and ability to handle sparse data (Chen & Guestrin, 2016).

It builds upon gradient boosting methods and has demonstrated superior performance in many data mining competitions and real-world applications. In the context of educational prediction, Márquez-Vera, Romero, and Ventura (2013) used genetic programming alongside XGBoost-like classifiers to predict student dropout and failure, achieving ROC AUC values between 0.65 and 0.75—comparable to the findings in the present study. These metrics suggestfair discriminatory power, particularly in imbalanced or multi-class educational datasets.

Furthermore, performance evaluation in academic prediction models extends beyond accuracy, as class imbalance often skews results. F1 score and ROC AUC are increasingly recognized as better indicators of a model's robustness (Huang & Fang, 2013). For example, a model may have high accuracy but low recall for underperforming students, thus failing in practical early warning systems. The moderate F1 score (0.407) and ROC AUC (0.642) in the current study mirror similar findings by Al-Barrak and Al-Razgan (2016), who used decision trees and reported modest predictive capacity for final GPA in Saudi universities.

In addition, despite the global growth in educational data mining, there remains a significant research gap in sub-Saharan Africa, where infrastructure, data availability, and model contextualization limit widespread adoption. Adejo and Connolly (2017) observed that predictive systems in Nigerian universities often lack integration into academic decision-making processes, partly due to data fragmentation and institutional resistance. Hence, implementing models like XGBoost in this context provides a pathway to not only predict outcomes but also identify key institutional and socio-economic variables affecting students' success.

Thus, while previous studies have successfully demonstrated the applicability of machine learning models in predicting academic performance, most are concentrated in developed countries with well-structured data systems. This raises concerns about the transferability of findings to low-resource settings. The current study addresses this gap by validating XGBoost within the Nigerian higher education system, thereby offering insights that can inform localized interventions and policy design.

3. METHODOLOGY

Here, the work examined research design, population and sample, data collection and variables, model development and ethical considerations. This study employed a quantitative, predictive research design using machine learning techniques to model and forecast undergraduate academic performance outcomes in Nigerian universities. The research aimed to evaluate the effectiveness of the XGBoost algorithm in predicting students' final degree classifications based on a structured dataset of academic, socio-economic, and institutional variables.

The population for this study comprised undergraduate students from selected Nigerian universities. A stratified random sampling technique was employed to ensure representation across faculties, years of study, and gender. The final dataset consisted of 1,200 undergraduate student records, collected over a three-year academic period from institutional databases and student records departments.

Data were obtained from official academic records, student registration portals, and structured institutional surveys. The dataset included both categorical and numerical variables. The dependent variable was academic performance, operationalized as final degree classification (First Class, Second Class Upper, Second Class Lower, Third Class). The independent variables included:

Student Factors: age, gender, attendance rate, study hours, internet access

Academic Factors: GPA in first and second years, continuous assessment scores, lecturer accessibility

Home and Socioeconomic Factors: parental education, occupation, family income, home location

Institutional Factors: school facilities, learning resources, class size, teaching methodology

Missing values were handled using mean/mode imputation for numerical/categorical data, and categorical variables were encoded using one-hot encoding.

Concerning model development, the dataset was split into a training set (80%) and a testing set (20%) using stratified sampling to preserve class distribution. The XGBoost classifier was implemented using Python's xgboost library. Model training involved hyperparameter tuning using a grid search with 5-fold cross-validation to optimize parameters such as:

Learning rate

Maximum tree depth

Subsample ratio

Number of estimators

The primary goal was to identify the best model configuration for predicting degree classification with acceptable generalization.

Model Evaluation Metrics

To assess model performance, the following evaluation metrics were used:

Accuracy: the proportion of correctly classified instances

F1 Score: harmonic mean of precision and recall, suitable for imbalanced class distributions

ROC AUC Score: area under the receiver operating characteristic curve, measuring the model's ability to distinguish between classes

Computation Time: time taken to train and validate the model

The final model achieved an accuracy of 0.430, an F1 score of 0.407, and a ROC AUC of 0.642, with a training time of 2.232 seconds.

As touching ethical considerations, ethical approval was obtained from the participating institutions' research and ethics committees. All student data were anonymized, and data usage adhered to institutional data protection and privacy policies. Informed consent was sought where surveys were used to augment institutional records.

4. RESULTS AND DISCUSSION

Table 1: *Accuracy of the validated model*

Model	**Accuracy**	**F1 Score**	**ROC AUC**	**Time (sec)**
XGBoost	0.430	0.407	0.642	2.232

The predictive model constructed with XGBoost classifier was tested on the test data with various metrics. The model gave an accuracy score of 0.430, F1 score of 0.407 and ROC AUC score of 0.642. The overall training time and validation time was 2.232 seconds. These findings indicate that although the model is a relatively efficient and effective way to compute, it has moderate predictive power.

The 43 percent rate of accuracy means that the model was right about the degree classification of students in fewer than half the instances. Although this accuracy can beat guessing at least in the context of a multi-class environment with four possible ways of classifying the degree this is not as accurate as what would be considered a high stakes academic decision-making standard. This conclusion is supported by the F1 score of 0.407, which shows that there is a moderate balance between precision and recall. This is especially significant with educational data, in which skew tends to favor middle level classifications (e.g., Second Class Upper/Lower), and may cause class imbalance (Huang & Fang, 2013).

The results were compared with works in the literature, and comparative observations showed that the findings were consistent with other earlier studies that had observed the same model results in learning institutions. As an example, Al-Barrak and Al-Razgan (2016) observed that both decision trees and ensemble models like XGBoost had similar accuracies (40 to 70 percent) to predict student GPA. Similarly, Márquez-Vera et al. (2013) reported that ROC AUC values in educational classification models tend to range from 0.6 to 0.75, particularly when dealing with high-dimensional and imbalanced data—supporting the ROC AUC of 0.642 reported in this study.

Moreover, the results reinforce the argument by Kotsiantis et al. (2004) that machine learning models in education often struggle to generalize well when datasets lack behavioral or psychological indicators (e.g., motivation, self-efficacy). The limited performance of the current model may reflect the need for broader and more diverse features beyond academic and demographic variables.

4.1 Predictive Modeling and Implications for Educational Policy in Nigeria

Within the Nigerian educational context, where student performance is influenced by complex socio-economic, infrastructural, and institutional factors, the moderate results of the XGBoost model underscore both the promise and limitations of predictive analytics. While the model's ROC AUC suggests it has some discriminative ability, the relatively low accuracy calls for cautious interpretation.

Factors such as poorly maintained institutional records, limited access to digital learning tools, and inconsistent curriculum delivery may have contributed to the model's limited performance. These challenges echo Adejo and Connolly's (2017) findings that Nigeria's higher education system lacks integrated data systems needed for effective learning analytics deployment.

The performance metrics indicate the need for future model improvement, which could include: Feature expansion (e.g., attendance logs, participation in tutorials, mental health indicators), data augmentation or resampling to address class imbalance, application of hybrid or deep learning models for better feature extraction, context-specific feature engineering to reflect Nigeria's educational dynamics. Future studies could also adopt longitudinal modeling techniques that incorporate time-series data to capture the progression of academic performance over semesters. Predictive modeling have the following implications on educational policy and practice in Nigeria

4.2 Early Intervention Strategies

Educational institutions can integrate machine learning models like XGBoost into their academic advising systems to identify at-risk students early. Even with moderate accuracy, such models can flag students who need academic support, mentoring, or counseling, especially when used in combination with human judgment.

4.3 Data-Driven Institutional Reform

The findings highlight the urgent need for universities to invest in centralized data systems that integrate academic, socio-economic, behavioral, and psychological variables. More comprehensive and structured data can significantly improve model performance and predictive insights.

4.3 Curriculum and Teaching Improvements

Academic performance predictions can inform reviews of teaching methodologies, particularly for courses with high failure rates. Insights from the model may point to curriculum bottlenecks or ineffective pedagogical practices that need re-evaluation.

4.5 Inclusive Educational Policy

Policy-makers can leverage predictive analytics to design targeted support programs—such as scholarships, digital resource access, and study skills workshops—for students from disadvantaged backgrounds, as socio-economic status emerged as a significant variable in many studies.

4.6 Capacity Building for Educational Data Mining

There is a pressing need to train institutional researchers and ICT staff in educational data mining techniques, data preprocessing, and ethical AI application to maximize the benefits of machine learning in Nigeria's educational landscape.

4.7 Recommendations

Based on the findings of this study and the identified limitations of the XGBoost model in predicting undergraduate academic performance, the following recommendations are proposed for researchers, academic institutions, and policymakers:

Expand the Range and Quality of Data Features: Academic institutions should prioritize the collection of comprehensive, high-quality datasets that go beyond academic records. Inclusion of behavioral, psychological, socio-cultural, and institutional variables (e.g., motivation, study habits, peer influence, access to digital tools, and school infrastructure) could enhance model performance and prediction reliability.

Address Class Imbalance through Data Preprocessing Techniques: Future research should apply data preprocessing methods such as SMOTE (Synthetic Minority Over-sampling Technique) or cost-sensitive learning to balance the class distribution and improve the F1 score and overall classification performance.

Explore Alternative and Hybrid Modeling Techniques: While XGBoost is powerful, researchers should explore other ensemble models (e.g., Random Forest, LightGBM), deep learning techniques, or hybrid models that may perform better on complex, non-linear educational data, particularly in multi-class prediction tasks.

Institutionalize Predictive Analytics in Higher Education: Universities should work towards integrating predictive models into student monitoring and advising systems, enabling data-driven decisions in identifying at-risk students and implementing timely academic interventions.

Build Capacity for Educational Data Science: There is a need to train faculty, administrators, and IT personnel in educational data mining, analytics, and ethical AI use. This capacity building will help institutions transition toward modern, technology-enhanced decision-making systems in line with global best practices.

Encourage Collaborative Research Across Institutions: Cross-institutional collaborations can help pool datasets from multiple universities, improving model generalizability and benchmarking performance across regions. National-level education bodies could facilitate such data sharing while ensuring privacy and ethical compliance.

Support Policy Frameworks for Learning Analytics: Government agencies and education regulators should develop national policies and ethical guidelines for the use of AI and predictive analytics in education. This includes data privacy laws, fairness protocols, and responsible use of student information.

5. CONCLUSION

This paper has discussed the relevance and effectiveness of the XGBoost machine learning algorithm in the determination of undergraduate academic outcomes, that is, final degree classifications in Nigerian universities. The findings showed that the model was characterized by a reasonable discriminative performance (ROC AUC = 0.642), moderate predictive accuracy (0.430), and F1 score (0.407), which means that the model has poor specificity in predicting academic performance.

These results are consistent with other studies of the difficulties of educational predictive modeling, especially in settings where data are imbalanced, features are limited, and where socio-institutional factors are complex. Though XGBoost demonstrated its promise, the medium performance of the model highlights the necessity of data quality, feature variability, and assimilation of the behavioral and contextual data of the students to enhance its performance. However, predictive machine learning use in academic prediction in Nigerian higher education is a significant move towards evidence-based decision-making. Despite its limitations, the model does include useful information about the trends in academic achievements and failures that can be used in institutional planning and early intervention initiatives.

Acknowledgments

None.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Funding Source

The authors received NO funding to conduct this study.

ORCID's

Moses Adekunle Ogunrinde [1] https://orcid.org/0009-0004-7852-9382
Eugenia A. Okwilagwe [2] https://orcid.org/0000-0002-8942-7435

REFERENCES

Adejo, O. W., & Connolly, T. (2017). Predicting student academic performance in higher education using data mining techniques. *Applied Computing and Informatics, 15*(2), 102–111. https://doi.org/10.1016/j.aci.2017.05.001

Al-Barrak, M. A., & Al-Razgan, M. (2016). Predicting students final GPA using decision trees: A case study. *International Journal of Information and Education Technology, 6*(7), 528–533. https://doi.org/10.7763/IJIET.2016.V6.745

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785

Huang, S., & Fang, N. (2013). Predicting student academic performance in an engineering dynamics course: A comparison of four types of predictive mathematical models. *Computers & Education, 61*, 133–145. https://doi.org/10.1016/j.compedu.2012.08.015

Kotsiantis, S. B., Pierrakeas, C., & Pintelas, P. (2004). Predicting students' performance in distance learning using machine learning techniques. *Applied Artificial Intelligence*, *18*(5), 411–426. https://doi.org/10.1080/08839510490442000

Márquez-Vera, C., Romero, C., & Ventura, S. (2013). Predicting school failure and dropout by using data mining techniques. *The Spanish Journal of Psychology, 16,* e42. https://doi.org/10.1017/sjp.2013.45

Nghe, N. T., Janecek, P., & Haddawy, P. (2007). A comparative analysis of techniques for predicting academic performance. Proceedings of the 2007 International Conference on Advanced Learning Technologies, 902–906. https://doi.org/10.1109/ICALT.2007.43

Nghe, N. T., Janecek, P., & Haddawy, P. (2007). A comparative analysis of techniques for predicting academic performance. *Proceedings of the 2007 International Conference on Advanced Learning Technologies*, 902–906. https://doi.org/10.1109/ICALT.2007.43