Fairness-Aware Predictive Analytics for Student Success in a Multicultural Higher Education Environment

Ganga Bhavani*, Afreen Arif, Ruhi Sethi
Amity Business School, Amity University Dubai, United Arab Emirates

HOW TO CITE:

Ganga Bhavani, Afreen Arif, Ruhi Sethi (2026). Fairness-Aware Predictive Analytics for Student Success in a Multicultural Higher Education Environment. International Journal of Special Education, 41(14s), 456-462

ABSTRACT:
This research will seek to create a fairness conscious predictive analytics framework to learn and facilitate student achievement in a diverse university environment. In transnational education context predictive models have been shown to be an effective way of to identify at-risk students, however, issues of algorithmic fairness and bias based on gender, nationality and scholarship status not yet been researched thoroughly. Anonymzed institutional data across multiple programs and semesters from a multicultural university context are used in this study. Predictive modeling is conducted using Logistic Regression and Gradient Boosting. Academic variables such as prior GPA and attendance are combined with demographic attributes, including age group, gender, nationality, and scholarship status. This study provides evidence of the importance of demographic variables relative to academic variables. This is critical for developing early warning systems, intervention strategies, and institutional policies in higher education organizations. The study is confined to four institutional setting and based on the accessible academic and demographic statistics. In order to maximize the strength, the data of four universities consisting of 16,500 students were involved in the analysis on the same set of features and modeling pipeline. University administrators and instructors to come up with data-driven and equitable intervention measures to address the needs of various learners can also use results. The fairness analysis framework has the potential to guide institutional policy about the responsible use of AI, early warning systems and specific academic support programs. This paper builds upon existing literature; it incorporates accuracy, interpretability, and equity in a single prediction analytics system designed to suit a transnational university environment.

Keywords: Predictive analytics; Student success; Fairness in AI; Algorithmic bias; Higher education; Interpretability; SHAP analysis; Gradient boosting; Logistic regression; Multicultural learning environment; Educational data mining

1.

2. Introduction

The role of machine learning models in higher education institutions in relation to predicting the performance of the students, their retention, and their progression are becoming increasingly important. When universities have access to reliable student data, these models can help identify students who may need additional academic or personal support. This allows institutions to take early action and create a learning environment that better supports student success. On a daily basis, there is an increased usage of using AI in decision-making in higher learning institutions, and predictive analytics stands out as the most effective tool.

In the context of a multicultural setting where education organizations face diverse student communities, the various demographic dimensions, which include gender and nationality, cannot be overlooked. In spite of the various predictive models, the heterogeneity of learning analytics, and the prominent role of algorithmic fairness in the predictive models of education, the area has not yet been extensively researched, particularly in foreign international branch campuses and thematic universities where cultural, social, and educational heterogeneity is prominent. This has given an opportunity to fill the existing gap through the support of literature and empirical analysis.

The demographic characteristics, such as gender, nationality, and socioeconomic background, in terms of their dimension, have been taken into consideration in the context of the analysis. In the past, very little research attention has been given to the transnational university environment, based on the perspective of algorithmic fairness in the framework of the multicultural higher education landscape. In this context, the specific research presents the involvement of empirical data that has received limited attention in the past.

1.1 Purpose and Objectives

The main research objectives of this research are to develop and test a fairness-oriented predictive analytics framework that can effectively predict student academic outcomes. The sub-objectives of the research are:

1. To predict the final course outcomes (Pass/Fail) based on demographic
2. To compare the performance of Logistic Regression and Gradient Boosting models
3. To identify the important determinants of students' success through SHAP-based interpretability techniques

Data used for the analysis consists of anonymized institutional data that spans across multiple programs and semesters in a multicultural university environment. Predictive modeling approaches that are used in this research are Logistic Regression Analysis and Gradient Boosting. In this research, academic predictor variables, such as GPA and attendance, are the rest of the paper is organized as follows: the next section reviews the literature relevant to the study, followed by methodology, and the final section is the discussion and conclusion. The conclusion of the paper will give the reader the overall findings and implications derived from it.

3. Review of Literature

2.1 Introduction to Learning Analytics in Higher Education

Learning analytics (LA) in higher education has evolved from describing what happened to predicting what might happen and recommending action. Descriptive analytics first examined data such as engagement and learning management system usage to identify past trends (Ngulube & Ncube, 2025), after which the field moved toward predictive analytics that uses machine learning to forecast outcomes such as dropout risk from academic and socioeconomic factors, enabling earlier intervention (Córdova-Esparza et al., 2025). The newest stage, prescriptive analytics, recommends specific actions for each student, such as adjusting course pacing or directing them to resources (Wang et al., 2024). Students generally prefer systems that give clear, detailed feedback on their progress (Corrin, 2021), yet this shift raises ethical concerns around privacy, consent, and transparency, as many are uncomfortable with the scope of data collection, making consultative policies essential (Corrin, 2021). Institutions must also avoid over-reliance on data that could disadvantage students with atypical learning paths or socioeconomic barriers (Rets et al., 2023). Ethical practice should include regular auditing and formal ethics roles (Nichols, 2024), transparency about data use, fairness checks across student groups (Rets et al., 2023), and compliance with legal requirements such as GDPR and FERPA (Córdova-Esparza et al., 2025); the growing use of process data from online assessments further increases the need for specialized ethics codes (Murchan & Siddiq, 2021).

2.2 Predictive Analytics for Student Success

Predictive models identify at-risk students by applying machine learning and statistics to varied data to forecast outcomes such as course failure, dropout, or retention, helping institutions intervene earlier. The typical workflow collects and cleans data, selects variables, trains algorithms, and tests predictions using standard evaluation measures, with data drawn from learning management systems, administrative records, and surveys. Datasets usually contain three groups of features. Academic and performance features such as grades, GPA, attendance, credits, midterms, and ongoing assessments are strong signals of progress (Al-Ahmad et al., 2025; Mehta & Pandey, 2025), and advanced methods such as Long Short-Term Memory networks or Support Vector Regression are sometimes used to predict GPA or exam results (Al-Ahmad et al., 2025; Aleena Fatima et al., 2025). Demographic and socioeconomic features add context but raise concerns about bias and fairness (Jawthari & Stoffa, 2022). Many algorithms build these systems: logistic regression remains popular for its interpretability and risk estimation (Nimy et al., 2023); tree-based methods such as decision trees and Random Forest handle mixed data and highlight important features (H. Elashmawi et al., 2025; Ridwan & Priyatno, 2024); Support Vector Machines suit high-dimensional data (Aleena Fatima et al., 2025); and boosting methods such as XGBoost perform well for dropout prediction with imbalanced data (Ridwan & Priyatno, 2024). Reliability is assessed through k-fold cross-validation and metrics such as accuracy, precision, recall, F1 score, and ROC AUC (Soni & Jain, 2024), alongside fairness tests for differential treatment of demographic groups (Jawthari & Stoffa, 2022). Once deployed, models trigger early alerts and targeted support such as mentoring and tutoring for high-risk students (Nimy et al., 2023; Soni & Jain, 2024).

2.3 Machine Learning Techniques Used in Education

Logistic Regression, Decision Trees, and Gradient Boosting are widely used in educational prediction, each balancing accuracy and interpretability differently. Logistic Regression is a simple linear classifier whose key advantage is interpretability: its coefficients show how strongly each feature relates to the outcome, helping educators see what drives student success or dropout risk (Eid & Raju, 2024). However, it cannot naturally capture nonlinear relationships or complex interactions, so it often underperforms more flexible models on accuracy and sensitivity (Xu, 2024). Gradient Boosting combines many weak models built sequentially, each correcting earlier errors, which yields stronger performance by learning complex nonlinear patterns; versions such as XGBoost and LightGBM achieve high precision, recall, and F1 scores in domains including credit risk, healthcare, and online-learning adaptability (Eid & Raju, 2024; Ngulube & Ncube, 2025; Xu, 2024). The tradeoff is reduced interpretability and greater computational and tuning demands.

2.4 Bias and Fairness in AI Systems

Algorithmic bias typically arises from the data, from model building and training, and from how results are interpreted and used. Bias often originates in training data: when groups are missing or distorted, the model learns patterns that fit them poorly, reducing accuracy and producing unfair decisions (Panić, 2025). It can also enter through proxy variables linked to protected traits such as race or gender, which can reproduce discrimination even when those traits are removed; resampling can rebalance data but may cause over fitting or reduced diversity (Panić, 2025). Model choices objective functions, evaluation metrics, feature selection, hyperparameter tuning, and decision thresholds—further shape outcomes, since optimizing only for overall accuracy can entrench inequalities for specific groups. Finally, bias appears in interpretation: failing to check whether a model makes different errors, such as more false negatives, for different groups can lead to unfair real-world decisions (Farayola et al., 2025), and a lack of transparency makes such bias harder to detect and correct (Sufian et al., 2024).

3. Methodology

3.1 Predictive Modeling

Two modeling models were used, namely logistic regression and gradient boosting (GBM). Logistic regression was chosen due to its interpretability and more probabilistic nature whereas GBM was added to determine the performance increase due to a more flexible and non-linear model. The models were trained on the training subset and tested based on the test subset on the basis of accuracy and precision, and recall among other common measures of performance.

3.2 Model Interpretation

The values of the shapley were calculated to determine the contribution of each of the features to the individual predictions, which would give a better understanding of the impact of what academic and demographic factors had the most significant contribution to the model. This approach allowed the clear interpretation and allowed the fairness to be studied by measuring the influence of attributes under protection on predictions.

4. Data Analysis

4.1 Dataset Overview

This research uses synthetically generated data to provide privacy reasons via institutional privacy limits as well as the sensitivity of individual-level student data as a result of the individual-level student record being sensitive in nature. In order to maximize the strength, the data of four universities consisting of 16,500 students. were involved in the analysis on the same set of features and modeling pipeline. This practice is in line with the best practices in ethical educational data mining and responsible AI research practices.

In this research, we have created simulated datasets of students (representing the demographic distribution and the academic profile of four multicultural universities in the UAE) in the academic year 456-462. The cohort sizes and sizes were 2,456-462,000 students to accurately represent the real enrollments of an institution.

4.2 Variables

The data set consists of three primary groups of variables namely, demographics, academic predictors and the outcome variable. Demographic variables are viewed as the attributes that are protected to analyze the fairness and they are gender (male/female), region of nationality (UAE, Other Arab, Asian, GCC, Europe, North America, Others), age (Below 20, 456-462, 456-462, 456-462, Above 35), scholarship status (Yes/No). These grouping variables were analyzed in order to identify possible biases in terms of gender, nationality and socioeconomic status.

Student engagement and student performance is captured in academic predictors, and it consists of such aspects as prior GPA (continuous, with a clipping of between 1.0 and 4.0), attendance rate (percentage), assessment score (percentage), course load (quantity of enrolled courses), and credit hours (12, 15, or 18). These characteristics were the ones that were supposed to be the most powerful predictors of course success because of their direct connection with academic performance.

The Final Outcome (Pass/Fail), outcome variable, was produced by a logistic model which dubbed GPA, attendance, assessment marks and course load together. The logistic transformation of a linear predictor was used to calculate probabilities and the results were finally sampled using a Bernoulli distribution which describes the chance of successful or unsuccessful course completion.

4.3 Model Performance

We have predictions from two models: a logistic regression and a GBM in Table 1.

Table 1: Accuracy from the Logistic regression and GBM.

Model	Accuracy
Logistic Regression	0.9963
Gradient Boosting Model	0.7421

Based on Table 1, we understand that the logistic regression model was performed exceptionally on the test data with an accuracy of 99.6, this means that it is virtually accurate in predicting Final Outcome. The Gradient Boosting Model (GBM) on the other hand achieved the accuracy of 74.2, which is significantly lower than that of the logistic regression. It means that in the case of this dataset, the simpler logistic regression model could be a better fit to consider the most important relationships among predictors and outcome, whereas the GBM could need more hyperparameter optimization or extra data preprocessing to perform better.

4.4. Shapley Analysis Summary

Table 2 is an analysis of the features contribution based on the shapley values. One feature that was analyzed and had a negative contribution to the predicted outcome was that of gender (male) with a ph = -0.00129; this means that the male gender would decrease the predicted outcome by a small margin. The highest positive contribution was on attendance rate (ph = 0.00124) implying that the higher the attendance the higher the chances of having positive outcome. Other characteristics such as previous GPA, nationality and age group had very minimal contributions thus suggesting that they do not have significant impact on the prediction. In general, feature contributions are very small, which is what is expected of a high probability that has already been very large (approximately 0.996); in that case, feature contributions are used to add value to the prediction but not to make it.

Table 2: Feature Contributions from the Shapley Values

Feature	Feature Value	Shapley Value (φ)	Variance (φ.var)
Gender	Male	-0.00129	9.37e-06
Nationality_Region	Asian	0.00069	2.84e-06
Age_Group	31_35	0.00015	1.40e-06
Scholarship_Status	No	0.000023	1.80e-07
Prior_GPA	2.5	0.00031	2.04e-06
Attendance_Rate	91.9	0.00124	9.16e-06

5. Conclusion

The paper created and tested a fairness-conscious prediction analytics system to explore student achievement in a multicultural higher education context. The integration of interpretable logistic regression with Gradient Boosting made the framework successful in predicting final course outcomes and also indicates the contribution of features and possible biases of the aforementioned attributes which are being protected. The logistic regression model was very accurate (99.6) that demonstrates the high predictive accuracy of the core academic and engagement variables that included the previous GPA, attendance percentage, and test scores.

Gradient Boosting, which is less precise (74.2%), provided a complementary view of non-linear effects of predictors.

Shapley analysis indicated that the individual features played a small role in predictions of this high-performing model, and the attendance rate demonstrated the greatest positive influence; gender had a small negative impact on the successfulness probability to be predicted. This proves that academic activity is the cause of the outcome, but, given the high predictive accuracy, demographic variables do not have a direct impact. Assessments of fairness have found gender and nationality differences where females had better true positive rates than males and certain nationalities e.g. Europe were not represented fairly in positive predictions. The status of scholarship had a slight bias, showing that the socioeconomic factors did not play a significant role in predictive disparities in this data. The conclusions help to focus on the fact that even the most accurate models can still reproduce the subtle biases that is why it is important to consider the metrics of fairness with the performance analysis.

Although the analysis used synthetic datasets that were calibrated with institutional data, the methodology can be applied to actual multicultural university environments, providing a feasible instrument of making data-driven decisions which are ethically responsible in the education sector.

References

Al-Ahmad, B. I., Alzaqebah, A., Alkhawaldeh, R., Al-Zoubi, A. M., Lo, H., & Ali, A. (2025). Predicting academic performance for students' university: Case study from Saint Cloud State University. *PeerJ Computer Science*, *11*, e3087. https://doi.org/10.7717/peerj-cs.3087

Aleena Fatima, Deebah Fatima, & Shurookh Mohammed Anana. (2025). Result Analysis and SGPA Prediction using SVR and LSR: A Comparative Analysis. *International Journal of Advanced Research in Science, Communication and Technology*, 577–589. https://doi.org/10.48175/IJARSCT-25176

Córdova-Esparza, D.-M., Terven, J., Romero-González, J.-A., Córdova-Esparza, K.-E., López-Martínez, R.-E., García-Ramírez, T., & Chaparro-Sánchez, R. (2025). Predicting and Preventing School Dropout with Business Intelligence: Insights from a Systematic Review. *Information*, *16*(4), 326. https://doi.org/10.3390/info16040326

Corrin, L. (2021). Shifting to digital: A policy perspective on 'Student perceptions of privacy principles for learning analytics' (Ifenthaler & Schumacher 2016). *Educational Technology Research and Development*, *69*(1), 353–356. https://doi.org/10.1007/s1456-462-09922-x

Eid, M. M., & Raju, S. K. (2024). Predicting Student Adaptability in Online Education: A Comparative Study of Machine Learning Models and Copula-Based Analysis. *Journal of Artificial Intelligence in Engineering Practice*, *1*(2), 62–71. https://doi.org/10.21608/jaiep.2024.318323.1008

Farayola, M. M., Tal, I., Saber, T., Connolly, R., & Bendechache, M. (2025). A fairness-focused approach to recidivism prediction: Implications for accuracy, trust, and equity. *AI & SOCIETY*. https://doi.org/10.1007/s0456-462-0456-462

H. Elashmawi, W., Rashad, M., & Alkady, Y. (2025). A Data-Driven Framework for Classifying Student Trajectories in Higher Education Using Machine Learning. *International Journal of Strategy and Organisational Learning*, *2*(1). https://doi.org/10.56830/IJSOL06202501

Jawthari, M., & Stoffa, V. (2022). Should Course-based At-risk Predication Models Include Protected Features? *2022 13th International Conference on Information and Communication Systems (ICICS)*, 257–261. https://doi.org/10.1109/ICICS55353.2022.9811127

Mehta, V., & Pandey, V. (2025). Predictive Analytics for Dropout Risk and Mental Fatigue Using Integrated Academic and Psychological Data Models. In Dr. S. Priya, Mr. R. H Gedam, Dr. V. Pandey, & Dr. D. K. Yadav (Eds.), *Digital Mental Health and AI in Campus Wellness Programs* (2025th ed., pp. 445–475). RADemics Research Institute. https://doi.org/10.71443/978934955456-462

Murchan, D., & Siddiq, F. (2021). A call to action: A systematic review of ethical and regulatory issues in using process data in educational assessment. *Large-Scale Assessments in Education*, *9*(1), 25. https://doi.org/10.1186/s4456-462-0456-462

Ngulube, P., & Ncube, M. M. (2025). Leveraging Learning Analytics to Improve the User Experience of Learning Management Systems in Higher Education Institutions. *Information*, *16*(5), 419. https://doi.org/10.3390/info16050419

Nichols, M. (2024). Development of an approved learning analytics ethics position. *Open Learning: The Journal of Open, Distance and e-Learning*, *39*(3), 212–225. https://doi.org/10.1080/02680513.2021.1986376

Nimy, E., Mosia, M., & Chibaya, C. (2023). Identifying At-Risk Students for Early Intervention—A Probabilistic Machine Learning Approach. *Applied Sciences*, *13*(6), 3869. https://doi.org/10.3390/app13063869

Panić, N. (2025). The impact of demographic composition in utKFace and appa-real datasets on fairness in age estimation models. *Journal of Process Management and New Technologies*, *13*(3–4), 78–96. https://doi.org/10.5937/jpmnt456-4620

Rets, I., Herodotou, C., & Gillespie, A. (2023). Six Practical Recommendations Enabling Ethical Use of Predictive Learning Analytics in Distance Education. *Journal of Learning Analytics*, *10*(1), 149–167. https://doi.org/10.18608/jla.2023.7743

Ridwan, A., & Priyatno, A. M. (2024). Predict Students' Dropout and Academic Success with XGBoost. *Journal of Education and Computer Applications*, *1*(2), 1–8. https://doi.org/10.69693/jeca.v1i2.13

Soni, M., & Jain, N. (2024). MACHINE LEARNING-BASED DROPOUT PREDICTION FOR UNDERGRADUATES. *ShodhKosh: Journal of Visual and Performing Arts*, *5*(5). https://doi.org/10.29121/shodhkosh.v5.i5.2024.4551

Sufian, M. A., Alsadder, L., Hamzi, W., Zaman, S., Sagar, A. S. M. S., & Hamzi, B. (2024). Mitigating Algorithmic Bias in AI-Driven Cardiovascular Imaging for Fairer Diagnostics. *Diagnostics*, *14*(23), 2675. https://doi.org/10.3390/diagnostics14232675

Wang, Y., Lai, Y., & Huang, X. (2024). Innovations in Online Learning Analytics: A Review of Recent Research and Emerging Trends. *IEEE Access*, *12*, 166761–166775. https://doi.org/10.1109/ACCESS.2024.3493621

Xu, T. (2024). Comparative Analysis of Machine Learning Algorithms for Consumer Credit Risk Assessment. *Transactions on Computer Science and Intelligent Systems Research*, *4*, 60–67. https://doi.org/10.62051/r1m3pg16