Available online at www.ijournalse.org

Emerging Science Journal

(ISSN: 2610-9182)

Vol. 9, Special Issue, 2025

" Emerging Trends, Challenges, and Innovative Practices in Education"

Predicting Dropout in MENA STEM Higher Education Using Explainable AI: A Machine Learning Approach

Reem A. M. Al Hashmi [1*], Panagiotis D. Zervopoulos [1], Hussein M. Elmehdi [2], Ilhan Ozturk [1, 3, 4]

[1] *College of Business Administration, University of Sharjah, Sharjah, P.O. Box 27272, United Arab Emirates.*

[2] *College of Sciences, University of Sharjah, Sharjah, P.O. Box 27272, United Arab Emirates.*

[3] *Faculty of Economics, Administrative and Social Sciences, Nisantasi University, Istanbul, Turkiye.*

[4] *University of VIZJA, Warsaw, Poland.*

Abstract

This study aims to develop an explainable machine learning–based early warning system to predict dropout risk among Science, Technology, Engineering, and Mathematics (STEM) students in the MENA region. Using longitudinal data from 6,798 undergraduate STEM students enrolled at a major UAE university, we evaluated six supervised classifiers: XGBoost, Gradient Boosting Machine (GBM), Random Forest, CART, Logistic Regression, and K-Nearest Neighbors. Models were trained on institutional student information system (SIS) data spanning ten cohorts (2010–2019), with class imbalance addressed through ROSE sampling. The top-performing models (XGBoost, GBM, and Random Forest) achieved AUC-ROC scores exceeding 0.91 and F1-scores above 0.84, significantly outperforming baseline models. Key predictors of dropout included the number of withdrawn semesters, second-term credit load, academic probation history, and performance in mathematics and physics. To improve interpretability, we applied SHapley Additive exPlanations (SHAP) analysis, enabling both global and individual-level feature attribution. The system offers scalable, real-time predictive capabilities using only routinely available SIS data, with no need for external surveys or learning management system inputs. The novelty of this research lies in its integration of explainable AI into a regional context, enabling early, transparent, and actionable interventions to reduce dropout. These findings contribute to data-driven retention strategies in higher education systems where predictive tools remain underutilized.

Keywords:

Machine Learning;
SHAP;
Explainable AI;
Student Dropout;
STEM Retention;
Educational Data Mining;
MENA.

Article History:

Received:	06	July	2025
Revised:	03	October	2025
Accepted:	14	October	2025
Published:	08	December	2025

1- Introduction

Student retention remains a persistent global challenge, particularly within Science, Technology, Engineering, and Mathematics (STEM) disciplines. Across many national contexts, STEM dropout rates exceed those of other academic fields, leading to institutional and societal consequences such as resource strain, lower rankings, and a diminished pipeline of skilled graduates vital to innovation and economic growth [1–4].

In the United Arab Emirates (UAE), STEM education plays a central role in national development strategies. The country has invested in curricular reform, international accreditation, and research infrastructure to support a knowledge-based economy [5, 6]. Yet, STEM attrition remains high, revealing a persistent gap between policy goals and student

* **CONTACT**: reemh@sharjah.ac.ae

DOI: http://dx.doi.org/10.28991/ESJ-2025-SIED1-016

outcomes [7, 8]. This is compounded by the absence of robust early intervention systems to identify and support at-risk students before failure or withdrawal occurs [9].

Traditional risk identification methods, such as GPA thresholds, probation, or reactive advising often miss complex, non-linear interactions among student, institutional, and contextual variables [10, 11]. In response, a growing body of research has adopted machine learning (ML) to develop proactive early warning systems using high-dimensional student data [12]. Techniques like decision trees, random forests, support vector machines, and neural networks have been used to predict dropout risk in diverse higher education settings, including STEM programs [13–15].

Despite their predictive strength, many ML models suffer from limited interpretability [16]. Without clear explanations for how predictions are generated, such models often face institutional resistance [17]. Recent research addresses this through explainable artificial intelligence (XAI) techniques, particularly SHAP (SHapley Additive Explanations) and LIME which enhance model transparency and support decision-making in higher education [18–21]. However, such methods remain largely absent in the Middle East and North Africa (MENA) region [22]. No studies to date have applied SHAP-based explainability to institutional Student Information System (SIS) data for dropout prediction, despite increased interest in data-driven policy and widespread collection of administrative records [23, 24]. Existing regional studies often rely on simple feature importance scores and lack the interpretability needed for ethical, actionable early warning systems [22].

This gap is especially relevant given the region's structural challenges. Policies related to scholarships, gender norms, and language of instruction introduce dropout risks not easily captured by Western-designed models [25, 26]. Furthermore, many MENA studies rely on static or single-term datasets, limiting their ability to track students longitudinally or confirm actual dropout outcomes [27].

This study addresses these limitations by using ten years of longitudinal SIS data (2010–2019) from a large UAE-based university to build a transparent, high-performing early warning system for dropout prediction. The dataset includes 6,798 STEM students and captures academic history, registration activity, and graduation status. Six ML algorithms are implemented, with SHAP used for both global and local interpretability. The results show that strong predictive performance can be achieved using routine academic data, such as GPA, probation status, and course withdrawal history without requiring external surveys or behavioral metrics. This enhances scalability and institutional feasibility across similar educational contexts.

The study is among the first in the MENA region to incorporate:

- Explainable AI (SHAP) for transparent decision-making at both institutional and individual levels;
- Longitudinal SIS data covering full student trajectories;
- Contextual framing aligned with policy, advising, and regional challenges.

Together, these contributions bridge the gap between predictive performance and implementation, offering a pathway for ethical and effective retention strategies. The study is guided by the following research questions:

- RQ1. Which ML model offers the best balance of accuracy and interpretability for dropout detection in the UAE?
- RQ2. Which demographic, academic, and institutional factors are the strongest predictors of attrition?
- RQ3. How can predictive analytics inform early intervention?

Conceptually, the study is grounded in Tinto's Student Integration Model, which emphasizes academic and social integration as central to student persistence. Within this framework, indicators such as low GPA, course withdrawals, and probation serve as measurable signs of disengagement [28, 29]. SHAP visualizations interpret these factors both globally (across all students) and locally (for individuals), enabling strategic and personalized interventions.

Finally, the study offers regionally grounded insights for policy and practice, including early risk detection in foundational courses and equitable support for international and multilingual students. These findings align with national goals in the UAE and broader MENA region to enhance retention and strengthen STEM education systems [30].

2- Literature Review

2-1- Theoretical Foundations: Student Dropout and Institutional Integration

Student attrition remains a persistent global challenge, especially in science and engineering disciplines where dropout rates are notably higher [31]. Tinto's Student Integration Model continues to be a foundational framework for understanding student persistence, conceptualizing dropout as a longitudinal process shaped by academic and social integration [28, 29]. Academic integration supports retention through GPA, engagement in coursework, and credit accumulation, while social integration reinforces institutional commitment by fostering peer networks, campus involvement, and a sense of belonging.

Empirical studies have operationalized these constructs using indicators such as academic probation, course withdrawals, and low cumulative GPA. Students exhibiting these patterns, particularly early in their academic trajectory, face elevated dropout risk [32, 33]. Socioeconomic pressures, including financial hardship and family obligations undermine academic and social engagement, compounding dropout risk [34].

Despite its conceptual strength, Tinto's model is often operationalized using linear statistical tools and fragmented data, limiting its predictive power [35]. These traditional approaches may overlook the nonlinear and heterogeneous nature of student pathways. In response, researchers have increasingly adopted machine learning (ML) methods capable of modeling multifactorial dropout risk [14]. When paired with explainable AI (XAI), these tools offer a scalable means to translate integration theory into actionable, transparent interventions for institutions [18].

2-2- Predictive Modeling of Dropout: From Statistics to Machine Learning

Statistical models such as logistic regression, survival analysis, and decision trees have long been used in higher education to identify dropout risk factors. These models offer interpretability and are well-suited for hypothesis testing, estimating dropout likelihood based on predictors like GPA, enrollment status, and demographics. However, they rely on assumptions of linearity and variable independence, which are often unmet in the complexity of student behavior. Their performance also declines in high-dimensional or imbalanced datasets and tends to overlook key interactions and subtle patterns [10, 11].

Machine learning (ML) models, such as Random Forest (RF), Gradient Boosting Machines (GBM), and eXtreme Gradient Boosting (XGBoost) have emerged as stronger alternatives. These algorithms detect nonlinear relationships, handle multicollinearity, and adapt to diverse data structures. Studies across Europe, North America, and Latin America show that ML consistently outperforms traditional methods in predicting dropout, especially in recall, precision, and early detection [12, 36, 37]. Their ability to process longitudinal and behavioral data has further cemented their role in early warning systems.

Yet, the rise of ML has raised concerns about transparency. These "black box" models often achieve accuracy without clarity, leaving stakeholders unsure of decision logic [22]. As a response, explainable AI (XAI) methods now help interpret model outputs, enhancing trust and aligning predictive tools with institutional accountability [18].

2-3- Explainable AI (XAI) and SHAP in Education

As machine learning becomes more integrated into educational systems, concerns about transparency and trust have intensified. Explainable Artificial Intelligence (XAI) addresses these concerns by making model predictions interpretable to non-technical stakeholders. Among the most widely used tools are SHapley Additive exPlanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME) [18, 38]. While both offer insights into how input features shape predictions, SHAP is particularly advantageous for its consistency and ability to explain predictions at both global (model-wide) and local (individual student) levels [20]. In high-stakes contexts like dropout prediction, such transparency is essential for institutional trust and actionable intervention [21].

Recent studies highlight SHAP's growing value in educational contexts. One study applied SHAP and other explainable AI techniques to predict academic performance, showing how key factors, such as academic progress and financial status influence student outcomes and dropout risk [38]. Another used SHAP in online learning environments to help advisors identify significant performance drivers, enabling more tailored support for individual learners [18]. These examples demonstrate how SHAP enhances interpretability, shifting from opaque risk scores to actionable insights. However, most XAI applications remain concentrated in Western or online contexts, where behavioral data are abundant. Few studies have applied SHAP to institutional SIS data in underrepresented contexts [22]. This creates a critical gap in deploying interpretable, SIS-based early warning systems in these contexts.

2-4- Dropout Determinants in the MENA Region

While global dropout models highlight academic preparedness, institutional integration, and financial aid, the MENA region presents distinct contextual factors. These include region-specific student funding policies, gender-based access constraints, and challenges related to multilingual instruction and curricular structure [25, 26]. For example, students in Gulf countries often receive full government-funded education and stipends but may still drop out due to rigid curricula, weak support in foundational courses, or academic probation policies. Gender-based dormitory access and parental restrictions can limit female students' university choices, an under-theorized factor in most dropout models [9, 39]. Likewise, the transition from Arabic-medium schooling to English- or French-medium instruction can impede academic engagement [25].

Despite these complexities, most MENA studies are descriptive or qualitative, focusing on performance, socioeconomic status, or motivation without analyzing large-scale institutional datasets [27]. A few recent studies from the UAE, Saudi Arabia, and Qatar have applied machine learning to longitudinal data (e.g., XGBoost, random forests, ensemble models) [23, 25, 40]. However, these models often function as "black boxes," lacking transparency or individualized insight for institutional planning [22].

2-5- Early Warning Systems and Institutional Integration

Early Warning Systems (EWS) in higher education are predictive tools designed to identify students at risk of failure or dropout before these outcomes occur. By tracking academic performance, engagement, and behavioral indicators, EWS enable institutions to intervene proactively through advising, tutoring, or outreach [12]. These systems operationalize Tinto's Student Integration Model by translating early signs of disengagement, such as declining GPA, repeated withdrawals, or low credit loads into alerts that can trigger timely support [33].

Globally, several universities have adopted ML-powered EWS to strengthen retention. In the UK, the Open University uses dashboards to monitor student activity and trigger personalized interventions [41]. In Spain, the Universitat Oberta de Catalunya supports online learners through predictive analytics [37]. Institutions in Latin America have similarly integrated ML models with behavioral and SIS data to issue early alerts [12]. These examples illustrate how predictive tools can enhance institutional workflows and student support services.

By contrast, such systems remain rare in the MENA region due to infrastructural barriers, lack of policy mandates, and concerns over the interpretability of ML models [27, 40]. While universities in the region often collect rich SIS data, they lack tools to translate it into actionable insights [23]. This prototype demonstrates the feasibility of explainable EWS design in underrepresented institutional settings.

3- Methodology

3-1- Machine Learning Models

This study is grounded in Tinto's Student Integration Model, which conceptualizes dropout as a consequence of weak academic and social integration. Accordingly, academic indicators such as GPA, course withdrawals, probation history, and earned credits were selected as model inputs and operationalized through machine learning (ML). The goal was to support interpretable risk prediction aligned with institutional priorities.

To ensure both predictive performance and interpretability, we employed six supervised learning algorithms. Random Forest was chosen for its ability to handle high-dimensional feature spaces and mitigate overfitting via its ensemble structure. Gradient Boosting offered iterative refinement by correcting prior classification errors, improving sensitivity to marginal cases. XGBoost was included as a fast, regularized variant of boosting, particularly effective for large, imbalanced datasets typical of educational systems. These ensemble models are widely validated in learning analytics for early risk detection, especially under class imbalance [13–15].

Complementing these, we included interpretable models to aid communication with institutional stakeholders. CART provides visual decision paths that are intuitive for non-technical audiences. Logistic Regression served as a conventional baseline, widely used in educational research. K-Nearest Neighbors, though more sensitive to feature scaling and structure, was included to offer a non-parametric contrast. Together, these models enabled a balanced assessment of accuracy, transparency, and institutional applicability [24, 32].

All models were implemented in R using randomForest, gbm, xgboost, rpart, class, and nnet. Data preprocessing was conducted with tidyverse, dplyr, and caret. Class imbalance was addressed using SMOTE and ROSE via the DMwR and ROSE packages [42, 43]. Model training was performed on the ROSE-balanced dataset with 5-fold cross-validation to reduce overfitting [20]. Default hyperparameters were retained across all models to ensure transparency, comparability, and ease of institutional implementation. Evaluation metrics and visualizations were generated using pROC, yardstick, and ggplot2, with all analyses conducted under a fixed seed (set.seed(123)). Figure 1 presents the three-phase pipeline: data preprocessing, model training and evaluation.

3-2- Model Training and Evaluation

Following class balancing using the ROSE technique, the dataset was randomly split into 70% training and 30% testing sets. All six machine learning models: Random Forest, Gradient Boosting Machine (GBM), XGBoost, CART, K-Nearest Neighbors (KNN), and Logistic Regression (LR) were trained on the balanced training data.

To improve generalizability and mitigate overfitting, five-fold cross-validation was applied during model training. This resampling approach ensured that performance estimates were stable and not overly dependent on any single subset of the data.

Model evaluation was conducted on the holdout test set using the following classification metrics:

- Accuracy: Overall correctness of the model's predictions.
- Precision: Proportion of predicted dropouts that were actual dropouts.
- Recall (Sensitivity): Proportion of actual dropouts correctly identified.
- Specificity: Ability to correctly classify students who did not drop out.
- F1-Score: Harmonic mean of precision and recall.
- AUC-ROC: Area under the Receiver Operating Characteristic curve.

Special emphasis was placed on F1-score and AUC-ROC as primary indicators of performance, given their robustness in imbalanced classification contexts. Together, these metrics offer a balanced assessment of each model's ability to detect at-risk students while minimizing false positives.

3-3- Feature Importance and Model Interpretability

To support model transparency and actionable insights, feature importance analysis was conducted on the three highest-performing classifiers: XGBoost, GBM, and Random Forest. These ensemble models were selected based on their strong predictive performance and established utility in educational data mining.

Each model's native gain-based ranking method was used to assess feature importance. Results highlighted the relative contribution of academic performance variables (e.g., GPA, earned credits) and behavioral indicators (e.g., course withdrawals, probations), consistent with prior literature on persistence and attrition.

To complement these global rankings with student-level interpretability, SHAP was applied to the XGBoost model. SHAP values quantify the marginal contribution of each feature to an individual prediction, capturing both the direction and magnitude of influence. Together, SHAP and gain-based methods offer complementary views (global and local) that enhance institutional understanding and ethical deployment of predictive analytics in student retention systems.

3-4- Ethical Considerations

This study was approved by the Research Ethics Committee at the "University Name Withheld for Review" (Reference Number: REC-25-03-12-02-PG, approval date: March 25, 2025). It involved the secondary analysis of anonymized administrative data obtained from the Office of the Registrar. No personally identifiable information was accessed, and there was no direct contact with students. All procedures complied with institutional ethical guidelines and data protection standards.

Access to SIS data and implementation of the early warning system prototype were facilitated by the first author's role as university registrar, which granted authorized administrative access. This positional alignment also ensured that model development remained grounded in institutional needs, supporting feasibility and potential integration into decision-making workflows.

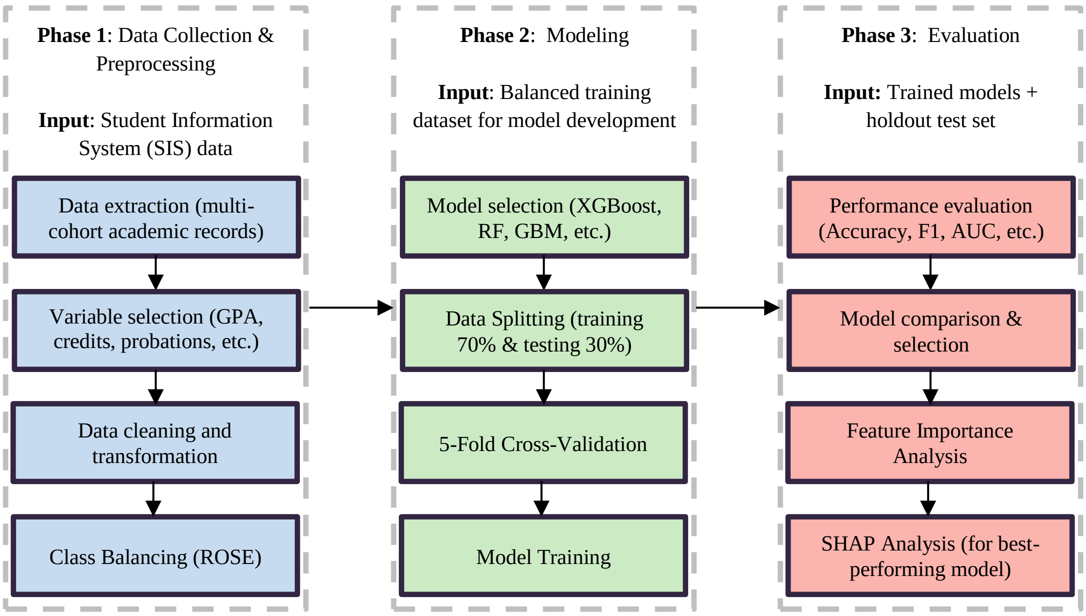

Figure 1. Workflow of the methodology showing the three-phase ML pipeline from SIS data preprocessing to model evaluation and explainability analysis

4- Sample and Data Description

This study applied a supervised machine learning framework to predict undergraduate dropout in STEM disciplines using institutional data from a large, research-intensive university in the UAE. The dataset comprised academic, demographic, and administrative records for all undergraduate STEM students across ten cohorts (2010–2019), extracted from the university's Banner Student Information System (SIS) via structured SQL queries.

A retrospective longitudinal design was employed, with the 2019 cohort serving as the endpoint to allow a minimum five-year observation window. This ensured conclusive classification of students as graduates (coded 0) or dropouts (coded 1), consistent with prior literature emphasizing extended observation to capture delayed attrition patterns [44]. The study's overarching aim was to move from retrospective reporting to real-time identification of at-risk students, forming the foundation for an early warning system.

Exclusion criteria included incomplete records, missing values, and transfer students who completed external STEM coursework, to ensure curricular consistency. Financial data were omitted due to the prevalence of institutional and external scholarships (~45%), which complicated interpretation. Disability status was excluded due to insufficient sample size.

Data preprocessing followed standardized and reproducible workflows in R. Categorical variables (e.g., FEMALE, UAE_NATIONAL, SCIENTIFIC_HIGHSCHOOL) were binary encoded; ordinal variables (e.g., PHYSICS_GRADE, MATH_GRADE) were mapped to the institutional GPA scale (0–4); and continuous features (e.g., GPA, earned hours, high school average, age at admission) were retained as numeric values. Feature names were aligned with SIS schema conventions and internally validated to ensure institutional interpretability and future integration into explainable AI dashboards.

The modeling design also reflects principles from student retention theory, particularly academic engagement and institutional integration, operationalized through quantifiable SIS features.

4-1- Sample Size and Variables

The final dataset included 6,798 undergraduate STEM student records. The binary outcome variable, DROP_OUT, classified students as either graduates (0) or dropouts (1). Sixteen independent variables were selected based on prior literature, expert input from registrar personnel, and their long-term availability in the SIS.

These variables were grouped into three categories:

- Demographic: gender, nationality, age at admission.
- High School Background: GPA and scientific-track status.
- University Academic Performance: term GPA, earned credit hours, grades in foundational STEM courses, and administrative indicators such as probation and withdrawal history.

This classification aligns with prior research. For example, demographic characteristics (e.g., gender, nationality) have been linked to persistence outcomes, particularly where social or financial constraints are present [40, 45]. Academic metrics such as GPA and credit hours frequently signal disengagement or academic difficulty [46, 47], while early struggles in core STEM subjects, especially physics and calculus often precede attrition [48, 49]. High school metrics offer a proxy for academic readiness [50], and institutional markers such as frequent withdrawals or incomplete grades are routinely flagged as concerns [39]. The final variable selection emphasized both theoretical relevance and operational feasibility. A full summary is presented in Table 1.

Table 1. Summary of Study Variables

Category	Variable Name	Description	Type
Demographic	FEMALE	Student's gender (1 = Female, 0 = Male)	Binary
	UAE_NATIONAL	Citizenship status (1 = UAE national, 0 = expatriate)	Binary
	AGE_AT_ADMISSION	Age at the time of university admission	Continuous
High School Background	HIGH_SCHOOL_AVERAGE	Final high school GPA (%)	Continuous
	SCIENTIFIC_HIGHSCHOOL	Whether the student came from a scientific track (1 = Yes, 0 = No)	Binary
Academic Performance	PHYSICS_GRADE	Performance in university physics (GPA scale: 0–4)	Ordinal
	MATH_GRADE	Performance in university math/calc (GPA scale: 0–4)	Ordinal
	CGPA_IN_FIRST_TERM	GPA after the first semester	Continuous
	EARNED_HOURS_IN_FIRST_TERM	Credit hours earned in the first semester	Continuous
	CGPA_IN_SECOND_TERM	GPA after the second semester	Continuous
	EARNED_HOURS_IN_SECOND_TERM	Credit hours earned in the second semester	Continuous
	NO_OF_PROBATIONS	Number of academic probations received	Integer
	NO_OF_FAILED_CRS	Number of failed courses	Integer
	NO_OF_WITHDRAWN_CRS	Number of withdrawn courses	Integer
	NO_OF_WITHDRAWN_SEMESTERS	Number of withdrawn semesters	Integer
	NO_OF_INCOMPLETE_GRADES	Number of incomplete (I) grades	Integer

Descriptive statistics indicated that 33.3% of students (n = 2,267) dropped out, while 66.7% graduated. The sample was 51.5% female and 32.6% UAE nationals, with an average high school GPA of 91.2% and a mean admission age of 17.7 years.

4-2- Addressing Class Imbalance

The dataset exhibited class imbalance: 66.7% of students graduated and 33.3% dropped out. This imbalance risks biasing machine learning models toward the majority class, reducing their sensitivity in detecting at-risk students. To address this, two resampling methods were evaluated: the Synthetic Minority Over-sampling Technique (SMOTE) and Random Over-Sampling Examples (ROSE). SMOTE, implemented via the DMwR package, generates synthetic dropout samples by interpolating between minority-class cases [42]. However, with default parameters (perc.over = 200, k = 5), SMOTE did not sufficiently rebalance the dataset. Future improvements may be possible through parameter tuning.

ROSE, implemented through the ROSE package [43] , applies a smoothed bootstrap strategy to oversample the minority class and undersample the majority. This approach produced a balanced dataset (50% dropout, 50% graduate) with improved class diversity, helping to mitigate overfitting.

Model evaluation favored ROSE, which yielded higher recall, F1-score, and AUC-ROC in top-performing models such as Random Forest, GBM, and XGBoost. These results underscore its effectiveness in enhancing generalization and improving early risk detection.

While synthetic balancing is common in predictive retention studies, it introduces artificial data. Future research may explore hybrid resampling or cost-sensitive methods to further strengthen fairness and model robustness.

5- Empirical Results

5-1- Descriptive Statistics and Dropout Trends

This section analyzes dropout patterns among 6,798 undergraduate STEM students, focusing on demographic and academic variables. The goal is to identify observable trends related to gender, nationality, and academic performance that can inform predictive modeling and institutional interventions.

5-1-1- Dropout Rates by Gender

Out of 6,798 students, 2,267 (33.3%) dropped out, while 4,531 (66.7%) graduated. Gender-based analysis showed a higher dropout rate among males (38.39%) compared to females (28.60%) (Table 2).

Table 2. Dropout Rates by Gender

Gender	Total Students	Dropouts	Dropout Rate (%)
Male (0)	3,295	1,265	38.39
Female (1)	3,503	1,002	28.60

Figure 2 illustrates this gap, with red indicating dropouts and blue representing graduates. The data suggest stronger persistence among female STEM students, potentially due to differences in motivation, academic habits, or support systems. Prior studies report that males in STEM face higher attrition risks, often linked to lower GPA, engagement issues, or inconsistent early performance [31].

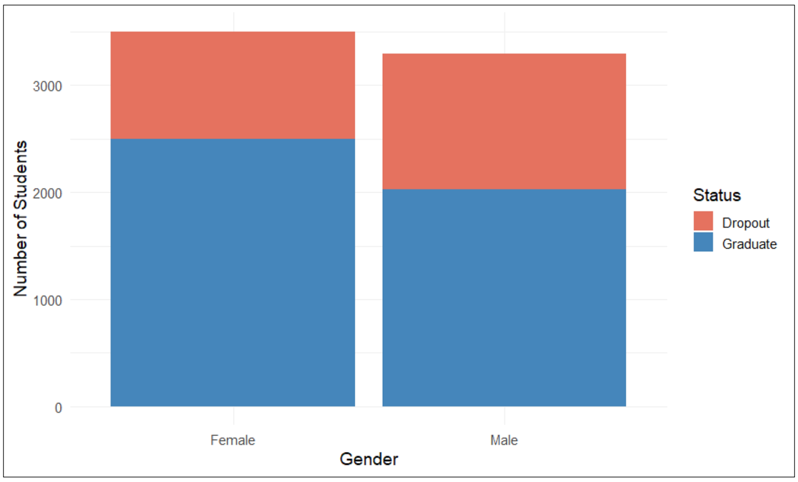

Figure 2. Dropout vs. Graduation by Gender

5-1-2- Dropout Rates by Nationality

Dropout patterns also varied by nationality. UAE nationals had a dropout rate of 30.97%, while expatriate students exhibited a slightly higher rate of 34.50% (Table 3).

Table 3. Dropout Rates by Nationality

Nationality	Total Students	Dropouts	Dropout Rate (%)
Non-UAE Nationals	4,583	1,581	34.50
UAE Nationals	2,215	686	30.97

Figure 3 visualizes these differences. Higher attrition among expatriates may stem from cultural adjustment challenges, limited support, and financial uncertainty [9].

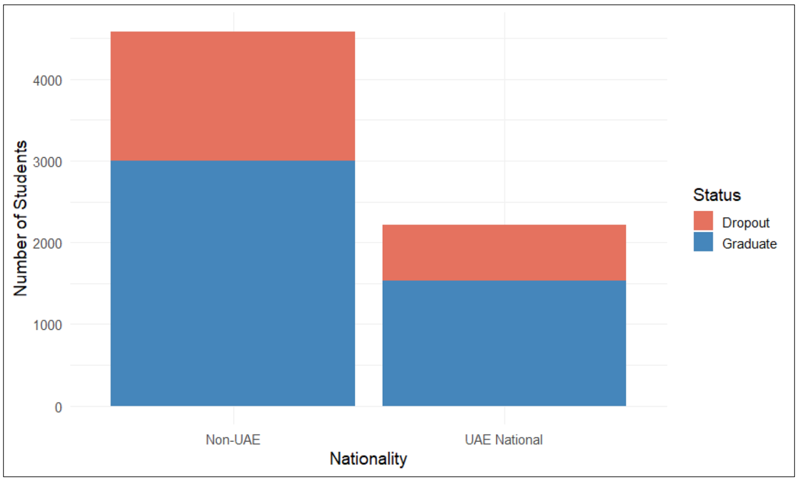

Figure 3. Dropout vs. Graduation by Nationality

5-1-3- Dropout Rates by First-Term GPA

First-term academic performance showed a strong link to dropout. Students with a GPA below 2.00 had a dropout rate of 62.02%, compared to 20.82% among those with a GPA of 3.00 or higher (Table 4).

Table 4. Dropout Rates by GPA Range

GPA Range	Total Students	Dropouts	Dropout Rate (%)
0.00 – 1.99	1,198	743	62.02
2.00 – 2.99	2,190	814	37.17
3.00 – 4.00	3,410	710	20.82

Figure 4 illustrates this trend. Early academic struggles may signal poor preparation or adjustment issues. Proactive support like tutoring and targeted advising can help mitigate attrition [35].

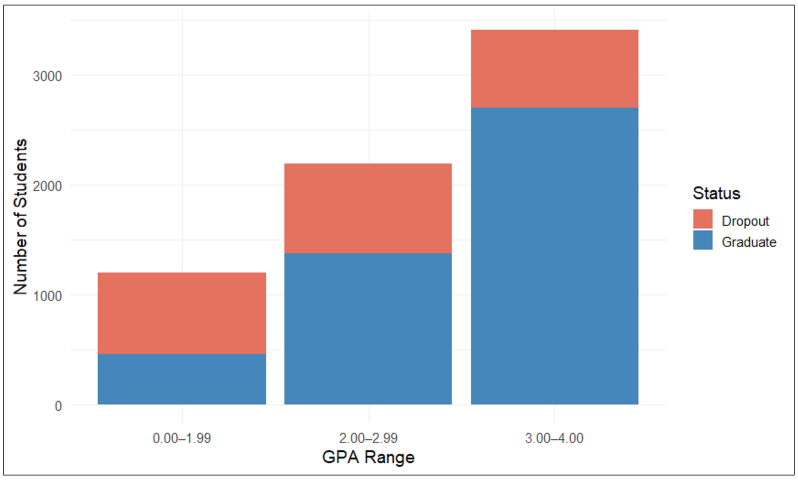

Figure 4. Dropout vs. Graduation by First-Term GPA

5-1-4- Dropout Trends by Gender and GPA

Table 5 shows dropout rates by gender and GPA. Among students with GPAs below 2.00, females had a higher dropout rate (66.85%) than males (54.86%). At higher GPA levels, males had slightly lower retention than females.

Table 5. Dropout Rates by Gender and GPA Range

Gender	GPA Range	Total Students	Dropouts	Dropout Rate (%)
Female	0.00 – 1.99	715	478	66.85
Female	2.00 – 2.99	1,137	463	40.72
Female	3.00 – 4.00	1,443	324	22.45
Male	0.00 – 1.99	483	265	54.86
Male	2.00 – 2.99	1,053	351	33.33
Male	3.00 – 4.00	1,967	386	19.62

Figure 5 reflects these trends. While GPA is a stronger predictor than gender alone, the interaction warrants further qualitative study [51].

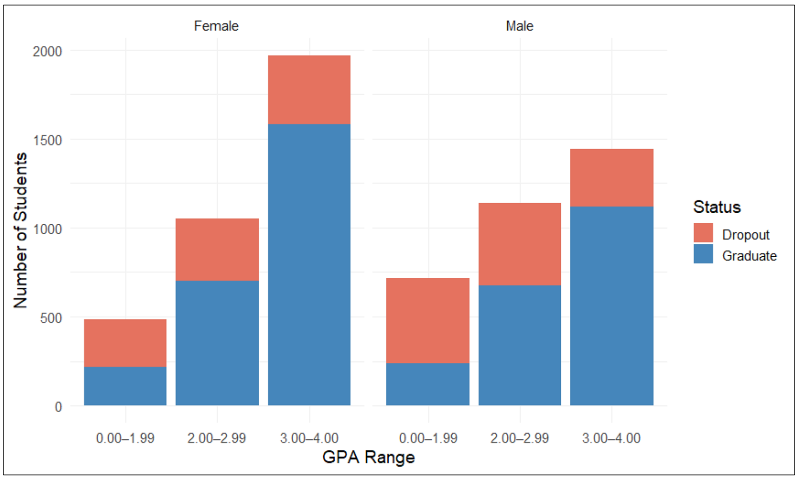

Figure 5. Dropout vs. Graduation by Gender and GPA

5-1-5- Descriptive Statistics Summary

Table 6 presents summary statistics. The mean first-term GPA was 2.84 (SD = 0.89), and the average admission age was 17.7 years. Most students had no academic probation (median = 0), but repeated probation appeared as a cumulative dropout risk.

Table 6. Descriptive Statistics of Key Variables

Metric	CGPA (First Term)	Age at Admission	Number of Probations
Mean	2.84	17.71	0.77
Median	3.00	17.00	0.00
Standard Deviation	0.89	1.11	1.54

5-2- Machine Learning Model Performance

Six evaluation metrics were used to assess the models' ability to predict dropout: accuracy, precision, recall, specificity, F1-score, and AUC-ROC. These metrics provide a comprehensive assessment of classification performance in the context of class imbalance.

Table 7 summarizes results for six classifiers: XGBoost, GBM, Random Forest, CART, Logistic Regression, and KNN. Ensemble models, especially XGBoost and GBM demonstrated the strongest performance, with AUC-ROC values above 0.91 and F1-scores exceeding 0.84. Their ability to balance recall and precision reflects their effectiveness in modeling complex feature interactions (Figure 6). Random Forest also yielded strong results, reinforcing the value of tree-based methods.

CART achieved the highest precision (0.8910), indicating strong identification of dropouts. However, its low recall (0.7119) suggests it missed many actual cases, which represents an important drawback for early warning systems, where recall is critical (Figure 7).

Logistic Regression and KNN, while interpretable and computationally simple, underperformed in recall and AUC-ROC. Their limitations in capturing nonlinear dropout patterns reduce their suitability for early risk detection.

Overall, XGBoost and GBM offered the best trade-off between sensitivity and specificity. Their robustness and high AUC values make them ideal candidates for institutional early alert frameworks that aim to identify at-risk students with high precision and minimal false negatives [13, 23].

Figure 8 illustrates the confusion matrices, further confirming the practical value of ensemble models. For example, XGBoost correctly identified 850 dropouts, misclassifying only 160, indicative of strong recall. In contrast, CART misclassified 940 dropouts, underscoring its high false negative rate despite precision. Likewise, KNN and Logistic Regression yielded excessive false negatives (890 and 797, respectively), limiting their practical application in real-time risk detection. The superior balance of precision, recall, and AUC-ROC among ensemble models supports their deployment in AI-powered early warning systems for timely student intervention [13, 23].

Table 7. **Model performance comparison across six classifiers**

Model	Accuracy	Precision	Recall	Specificity	F1-Score	AUC-ROC
XGBoost	0.8503	0.8543	0.8416	0.8589	0.8479	0.9162
GBM	0.8420	0.8517	0.8248	0.8589	0.8380	0.9144
Random Forest	0.8469	0.8490	0.8406	0.8531	0.8448	0.9136
CART	0.8140	0.8910	0.7119	0.9144	0.7914	0.8161
K-Nearest Neighbors	0.7242	0.8094	0.5802	0.8658	0.6759	0.7879
Logistic Regression	0.7282	0.7484	0.6802	0.7753	0.7127	0.7846

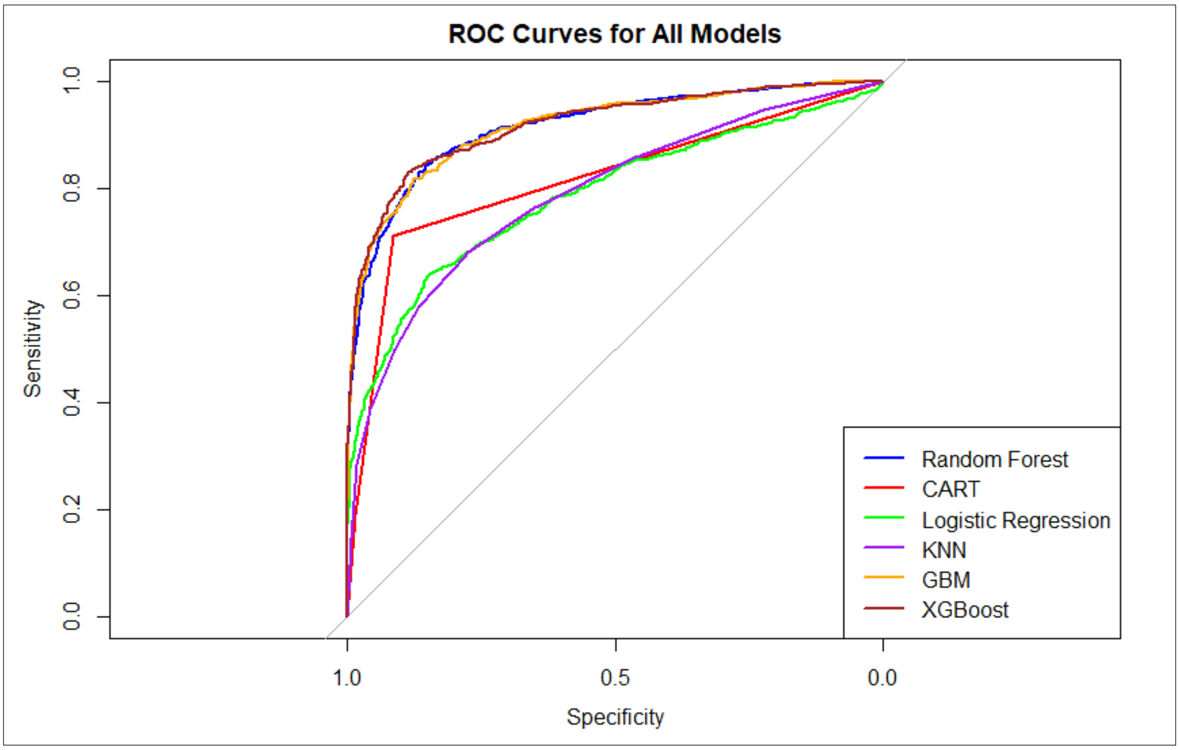

Figure 6. **ROC curves comparing dropout prediction performance across six classifiers**

Receiver Operating Characteristic (ROC) curves show that XGBoost, GBM, and Random Forest achieved the highest AUC values, indicating strong overall classification ability. This suggests these models are effective at distinguishing dropouts from non-dropouts, even under data imbalance.

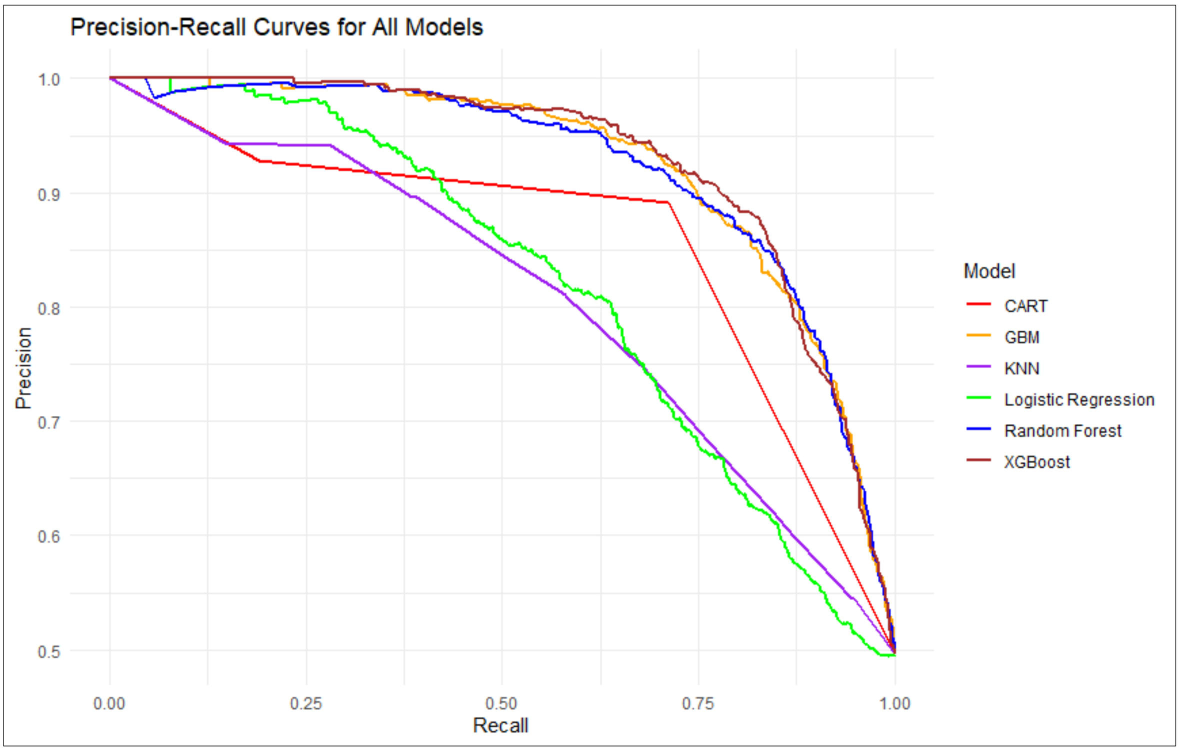

Figure 7. Precision–Recall curves for six classifiers under dropout class imbalance

The PR curves show that XGBoost and GBM maintain a superior balance between precision and recall, making them suitable for early intervention strategies. CART's high precision but low recall reveals a trade-off that may lead to missed dropout cases.

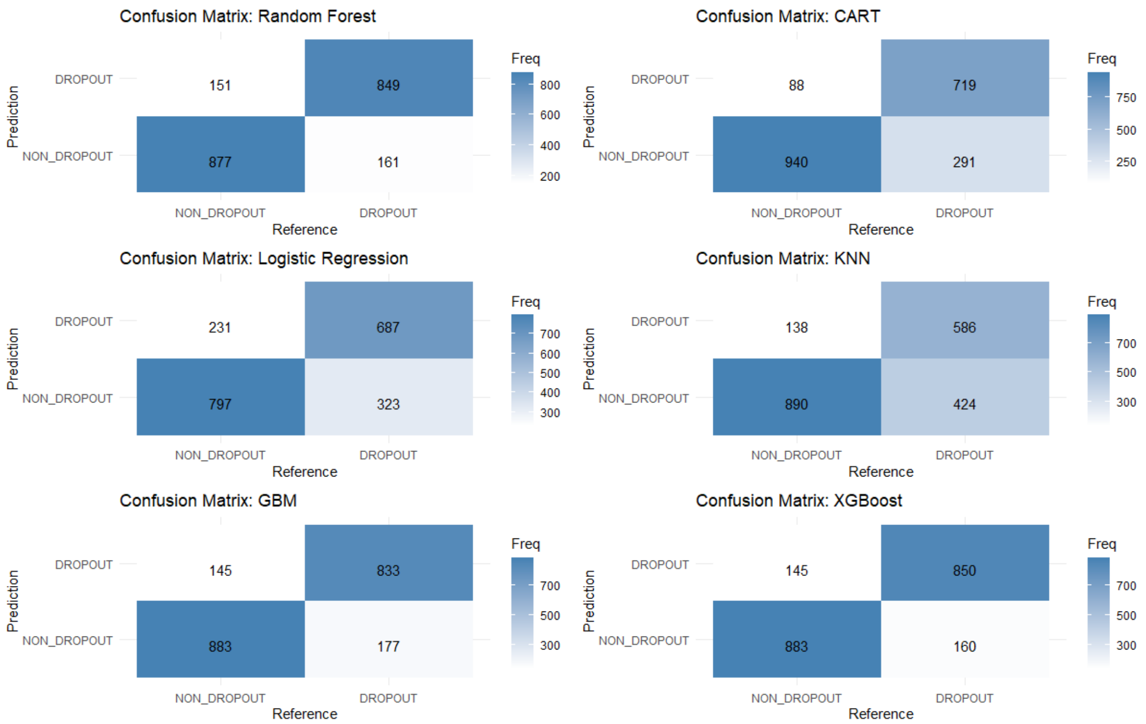

Figure 8. Confusion matrices for six classifiers showing prediction accuracy across dropout categories

Each matrix displays the number of correctly and incorrectly classified instances for dropout and non-dropout students. The diagonal elements represent accurate predictions (true positives and true negatives), while off-diagonal values reflect misclassifications. Ensemble models (XGBoost, GBM, Random Forest) demonstrated stronger classification balance by minimizing false negatives (dropouts misclassified as non-dropouts) and maximizing true positives, both of which are critical for early intervention.

5-3- Model Interpretability and Feature Importance

To support explainability and intervention design, feature importance was analyzed using the top-performing models: XGBoost, GBM, and Random Forest. Each model achieved an AUC-ROC above 0.91 and an F1-score above 0.84 (Table 7). These models were selected for their consistent performance and reliability in identifying early risk predictors.

5-3-1- Feature Ranking Across Models

Bar plots were used to visualize the top ten features for each model (Figure 9). The most influential predictors were consistent across models, reinforcing their reliability for early risk detection.

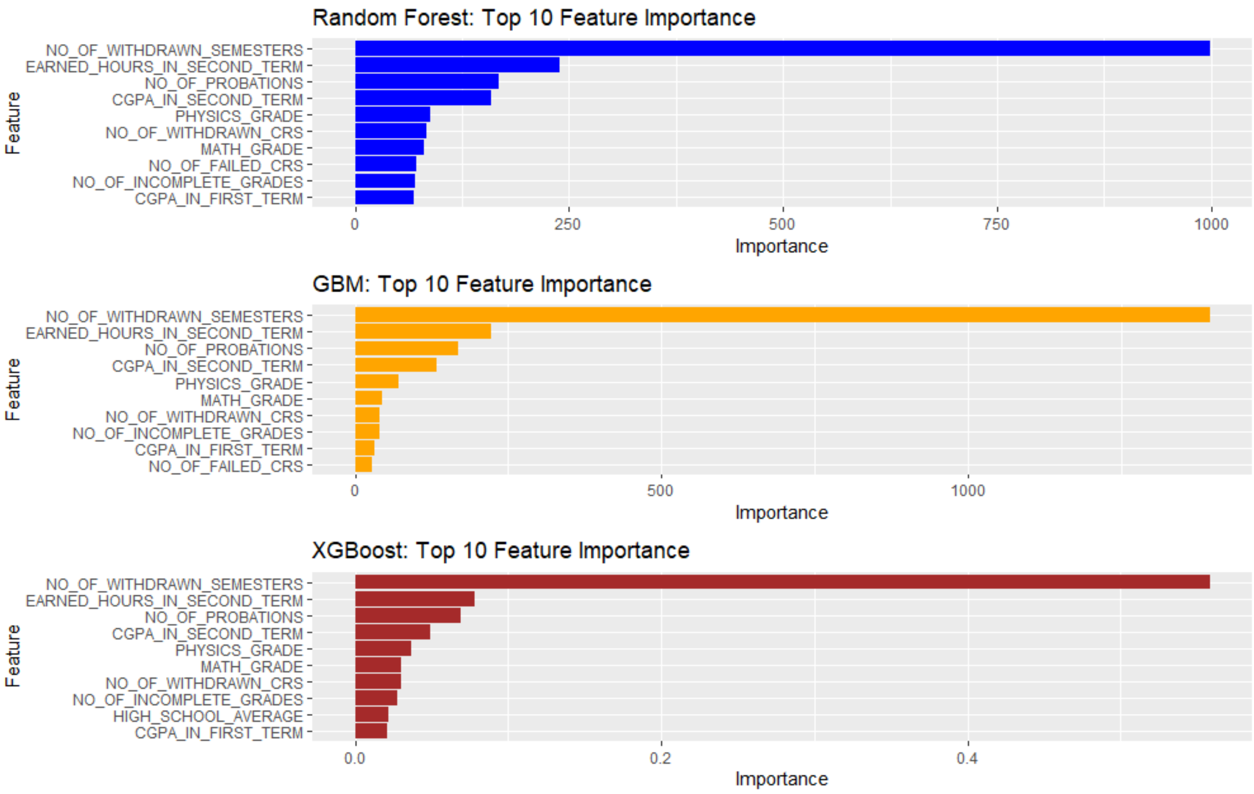

Figure 9. Top ten predictive features for dropout across ensemble classifiers

Bar plots display the top ten features influencing dropout prediction in the Random Forest, GBM, and XGBoost models. Academic indicators (e.g., CGPA, course grades, credit hours) and behavioral variables (e.g., withdrawn semesters, probations) consistently emerged as key predictors. Feature definitions are provided in Table 8.

Table 8. Description of Features with the Highest Importance Across XGBoost, GBM, and Random Forest

Rank	Feature	Description
1	NO_OF_WITHDRAWN_SEMESTERS	Total semesters withdrawn during enrollment
2	EARNED_HOURS_IN_SECOND_TERM	Credits completed in the second semester
3	NO_OF_PROBATIONS	Number of academic probations
4	CGPA_IN_SECOND_TERM	GPA after second semester
5	PHYSICS_GRADE	Grade in university physics
6	MATH_GRADE	Grade in university math/calc

The prominence of withdrawn semesters, second-term credit hours, and academic probations underscores the importance of maintaining academic continuity and momentum. These variables often serve as early warning signals, identifiable before CGPA declines to critical levels. Notably, EARNED_HOURS_IN_SECOND_TERM emerged as a more influential predictor than first-term GPA in some models, suggesting that academic behavior tends to stabilize or deteriorate more visibly after the initial transition period. This aligns with prior findings indicating that second-term performance represents a critical inflection point in student persistence [13].

Course-specific grades in physics and math also proved more predictive than cumulative GPA in certain models. This result may reflect the gatekeeping function of foundational STEM courses in shaping student confidence and progression through degree requirements. For example, failure in a core subject such as physics can trigger cascading disengagement, particularly in tightly sequenced STEM programs.

These insights offer a more granular understanding than traditional metrics, which typically emphasize cumulative GPA. SHAP analysis (see Section 5-3-2) further contextualizes these findings by illustrating how specific predictors influence dropout risk at the individual level.

5-3-2- Interpretability Through SHAP Analysis

To enhance transparency and institutional adoption, a SHAP summary plot was generated for the final XGBoost model (Figure 10). SHAP values quantify both the magnitude and direction of each feature's contribution to predictions, enabling interpretation at the individual student level. As shown in Figure 10, the most influential feature was the number of withdrawn semesters (mean SHAP = 1.828). Higher values (purple dots on the right) consistently yielded positive SHAP values, indicating that withdrawal history significantly increased dropout risk. This aligns with disengagement theories and flags withdrawal frequency as an actionable risk factor.

Second-term academic performance was also critical. CGPA_IN_SECOND_TERM (mean SHAP = 0.225) and EARNED_HOURS_IN_SECOND_TERM (0.373) showed that lower values (yellow dots on the right) increased dropout risk. While CGPA_IN_FIRST_TERM also contributed (SHAP = 0.203), second-term indicators had greater influence, suggesting that continued underperformance outweighs initial struggles.

Grades in key STEM courses: MATH_GRADE and PHYSICS_GRADE (SHAP = 0.255 and 0.291), highlighted the gatekeeping role of foundational courses. Students with higher grades (purple dots on the left) had lower dropout risk, underscoring the importance of targeted academic support in these subjects. The number of academic probations (NO_OF_PROBATIONS, SHAP = 0.400) was often more predictive than GPA, suggesting that probation status may reflect both academic and psychological risk.

Features such as HIGH_SCHOOL_AVERAGE and NO_OF_INCOMPLETE_GRADES had lower SHAP values but contributed meaningfully in specific cases. SHAP results confirmed gain-based rankings (Table 8) and provided fine-grained insight for identifying students at risk before CGPA thresholds are breached. Although CGPA_IN_SECOND_TERM had moderate multicollinearity with CGPA_IN_FIRST_TERM (VIF = 6.01), its SHAP impact was distinct and meaningful. Tree-based models like XGBoost manage multicollinearity effectively, and second-term GPA captures academic trends beyond initial performance.

For example, one student with a CGPA of 2.16, three failed courses, and 0.12 withdrawn semesters had a high predicted dropout probability. SHAP identified withdrawn semesters (SHAP = –1.46), low physics grade (3.22; SHAP = –0.30), and limited earned hours as dominant risk factors. Although second-term GPA (SHAP = +0.19) and earned hours (SHAP = +0.33) mitigated risk slightly, they were insufficient. This illustrates how SHAP can guide early advising by pinpointing the most influential dropout risks.

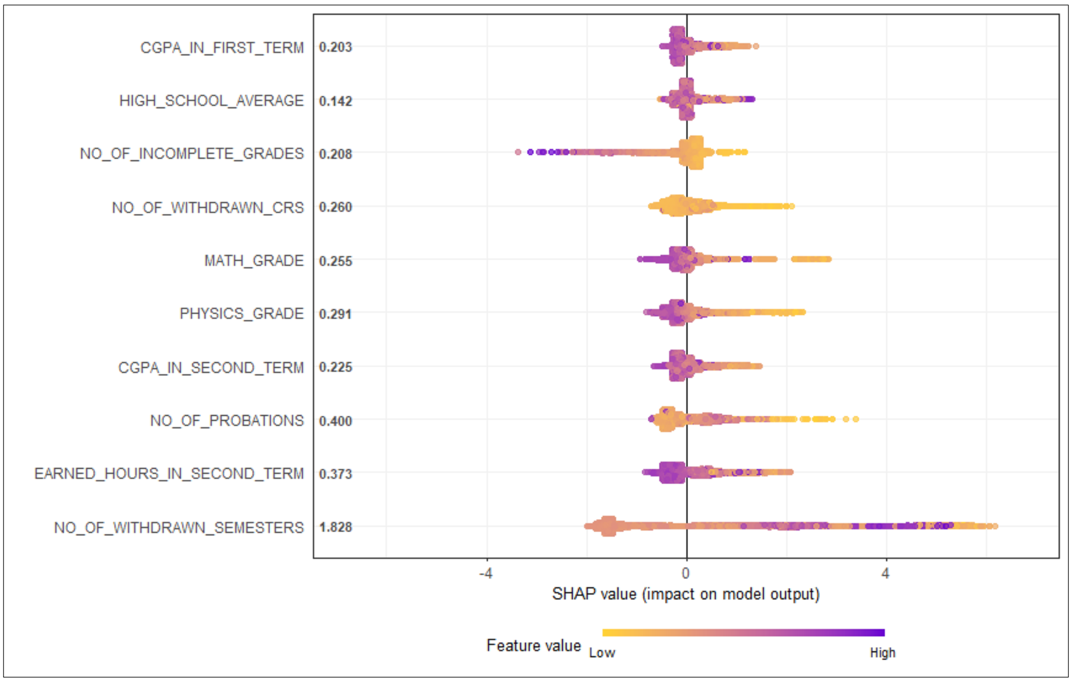

Figure 10. SHAP summary plot for the XGBoost model illustrating dropout prediction features

The plot ranks predictors by mean absolute SHAP value, showing each feature's impact on model output. Dots represent students; color indicates feature value (purple = high, yellow = low), and position reflects dropout risk contribution. Withdrawn semesters, probation history, and second-term performance had the greatest influence. The plot aids localized interpretation and informs early intervention.

6- Discussion

6-1- Interpretation of Findings and Theoretical Implications

This study developed a machine learning-based early warning system to predict dropout among STEM students at a major UAE university. Using longitudinal data from ten cohorts (2010–2019), it addressed a research gap in the MENA region by leveraging academic and demographic variables from the student information system (SIS). The analysis confirmed that withdrawn semesters, second-term credits, probation history, CGPA, and performance in mathematics and physics were the most influential predictors across top-performing models (XGBoost, GBM, Random Forest).

These findings align with international research while offering regional insights. The study operationalizes Tinto's Student Integration Model [28, 29], treating academic indicators such as course withdrawals and probation as proxies for disengagement.

By embedding these into predictive models, the study moves Tinto's framework from theory to application, enabling real-time institutional response. This aligns with current directions in educational data mining, which emphasize predictive, theory-informed systems [52, 53]. SHAP-based interpretability further enhances alignment between model outputs and institutional support strategies.

In addition, this study offers one of the first documented applications of SHAP-based explainable AI to SIS-integrated dropout prediction in the MENA region. Unlike prior models that treat interpretability as a secondary concern, explainability here was embedded from the outset to support academic advising in a practical, actionable manner.

6-2- Comparison with Global Literature

Gender-based dropout patterns in this study diverge from those typically reported in Western literature. In the U.S. and Europe, women are underrepresented in STEM fields and face higher dropout risks due to stereotype threat, limited belonging, and male-dominated academic climates. However, once enrolled, women often show comparable or stronger persistence in early years, with challenges emerging later, particularly in on-time graduation and confidence-related barriers [54, 55].

In contrast, our UAE-based analysis, consistent with broader MENA findings revealed higher dropout rates among male students. Regional studies attribute this to labor market incentives that encourage early employment, especially in the public sector, and to weaker academic engagement among men. For women, higher persistence is linked to stronger study habits, family support, and cultural framing of university as a safe or socially endorsed space. While some female dropout cases are influenced by marriage or family obligations, these are often temporary or context-specific rather than systemic [9, 31].

Academic probation and course withdrawal also carry distinct cultural meanings. In the MENA context, these are often perceived as failures that generate shame and familial disappointment, discouraging help-seeking [9]. Conversely, in U.S. and European institutions, while stigma exists, probation is more frequently framed as a recovery mechanism within a growth-oriented academic narrative [56].

Socioeconomic disparities offer another point of divergence. In Western contexts, lower SES is a consistent dropout predictor [33, 44], while in the GCC, UAE nationals—often from lower-SES households—exhibit higher persistence than academically stronger expatriates. This reversal is explained by state-sponsored education, stipends, and guaranteed employment pathways, which buffer nationals against dropout risks. Expatriates, by contrast, face visa-linked consequences and limited institutional safety nets [9, 25].

Together, these findings highlight the limitations of directly applying Western dropout models in the Gulf and underscore the importance of culturally grounded early warning systems and support strategies.

6-3- Implications and Institutional Recommendations

This study offers a set of practical recommendations for institutional policy and student retention strategies grounded in explainable AI. The strong performance of XGBoost, GBM, and Random Forest models (AUC-ROC > 0.91; F1 > 0.84) supports the feasibility of implementing machine learning–based early warning systems across higher education institutions. Importantly, these models rely on student information system (SIS) data already available at most universities, enhancing scalability and integration potential.

Universities should consider deploying automated risk classification tools that flag students based on critical indicators such as withdrawn semesters, credit accumulation by the second term, probation history, and GPA. When a predefined risk threshold is met, the system can notify academic advisors and trigger proactive outreach.

Recommended interventions include:

- Mandatory academic counseling for students experiencing repeated course withdrawals, multiple probations, or delayed credit progression;
- Early academic support in gateway STEM subjects, particularly mathematics and physics, where poor performance is an early predictor of dropout;
- Tailored advising frameworks for expatriate students, who may face heightened financial, legal, or cultural stressors.

These targeted interventions should be embedded within a broader adaptive support infrastructure. Core components of such a system include:

- Continuous risk monitoring using real-time SIS updates;
- Role-specific dashboards for faculty, advisors, and support staff;
- Automatic referrals to tutoring, financial aid, or mental health services;
- Dynamic feedback loops that revise risk status as new academic data are ingested.

At scale, the proposed models could identify up to 83.3% of eventual dropouts by the end of the second semester, providing institutions a valuable window for timely, data-informed intervention.

Crucially, the findings reinforce the importance of equity-oriented, culturally responsive practices. Male students and non-UAE nationals demonstrated higher risk levels, indicating the need for gender-sensitive advising, national identity–aware support structures, and AI systems that account for stigma and social pressures. These recommendations align with emerging calls for fairness-aware, transparent, and contextually grounded uses of AI in educational environments [57].

6-4- Limitations and Future Research

This study presents several limitations that also offer directions for future inquiry. First, the predictive models were built exclusively using academic and demographic data extracted from the student information system (SIS). Consequently, critical dimensions such as student engagement, psychological well-being, and financial strain were not included due to institutional data constraints. Future research should incorporate behavioral signals from learning management systems (LMS), counseling services, and financial aid records to develop more comprehensive and context-sensitive risk profiles.

Second, the analysis focused on STEM students at a single UAE university. While this context-specific scope enhances internal validity, it limits generalizability. Institutional structures, advising practices, and cultural norms differ across MENA universities. Expanding future studies to include multi-institutional datasets and non-STEM disciplines would strengthen external validity and support scalable early warning system design.

Third, although ensemble models such as XGBoost, GBM, and Random Forest demonstrated high predictive performance, their algorithmic complexity challenges interpretability. SHAP (Shapley Additive Explanations) was applied to improve model transparency, but further research is needed to operationalize explainable AI (XAI) in institutional settings. Emerging work [58] shows SHAP's potential to uncover behavioral risk patterns even in blended or sparse-data environments.

Fourth, while this study employed conventional supervised learning and a standard train-test validation pipeline, future research could experiment with advanced methodologies such as longitudinal modeling, cost-sensitive learning, or ensemble stacking. These approaches may better reflect the temporal dynamics and fairness considerations inherent in dropout prediction, though practical tradeoffs in complexity and implementation must be carefully assessed.

Fifth, the long-term institutional impact of AI-enabled early warning systems remains underexplored. Future studies should evaluate how predictive models influence advising practices, student trajectories, and institutional policy, while also addressing ethical, equity, and trust dimensions in live deployment.

Lastly, although data balancing was achieved using SMOTE and ROSE, more sophisticated techniques such as SMOTE-ENN or hybrid resampling were not applied. Future work could investigate these alternatives to determine whether performance gains justify potential losses in transparency and usability.

7- Conclusion

This study developed and evaluated a machine learning–based early warning system to predict undergraduate dropout in STEM disciplines using institutional data from a major UAE university. Drawing on longitudinal SIS records from 6,798 students across ten cohorts (2010–2019), the analysis identified key predictors of attrition, including withdrawn semesters, academic probation, low second-term credit loads, and weak performance in foundational courses; particularly mathematics and physics. Among six models assessed, ensemble methods such as XGBoost, Gradient Boosting Machine (GBM), and Random Forest demonstrated superior predictive accuracy (AUC-ROC > 0.91), significantly outperforming traditional models such as logistic regression and KNN. These findings affirm the practical utility of institutional SIS data and support the deployment of advanced machine learning for early dropout risk detection.

The study offers both theoretical and applied contributions. The strong influence of early academic momentum reinforces the relevance of Tinto's Student Integration Model, while the use of SHAP-based model interpretability enhances the transparency and actionability of results. Furthermore, the elevated dropout risk observed among male and non-UAE national students highlights the need for targeted, equity-oriented interventions. Although the scope was limited to academic and demographic data from a single institution, the predictive framework is adaptable and scalable. Future research should integrate behavioral, psychological, and financial variables to improve accuracy and explore generalizability across other disciplines and institutional contexts. Ultimately, this research presents a replicable, explainable, and context-sensitive approach to student retention. By embedding AI-driven early warning tools into academic advising systems, universities can transition from reactive responses to proactive support—reducing inequality, improving degree completion, and promoting institutional accountability through data-informed intervention.

8- Declarations

8-1- Author Contributions

Conceptualization, R.A.M.A.H., P.D.Z., and H.M.E.; methodology, R.A.M.A.H. and P.D.Z.; software, R.A.M.A.H. and P.D.Z.; validation, R.A.M.A.H. and P.D.Z.; formal analysis, R.A.M.A.H. and P.D.Z.; investigation, R.A.M.A.H.; data curation, R.A.M.A.H.; writing—original draft preparation, R.A.M.A.H.; writing—review and editing, P.D.Z. and I.O.; visualization, R.A.M.A.H.; supervision, P.D.Z., H.M.E., and I.O. All authors have read and agreed to the published version of the manuscript.

8-2- Data Availability Statement

The data presented in this study are available on request from the corresponding author. The data are not publicly available due to institutional restrictions and confidentiality agreements, as they involve sensitive student information.

8-3- Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

8-4- Acknowledgements

The authors thank Mr. Shehab Fakhry Abushehab, Senior Registration System Officer at the University of Sharjah, for his technical support in developing SQL functions that enabled the extraction of key variables from the student information system.

8-5- Institutional Review Board Statement

The study was approved by the Research Ethics Committee at the University of Sharjah (Reference Number: REC-25-03-12-02-PG, approval date: March 25, 2025). Ethical approval was granted for the secondary analysis of anonymized institutional data, with no access to personally identifiable information and no direct interaction with students.

8-6- Informed Consent Statement

Not applicable

8-7- Conflicts of Interest

The authors declare that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

9- References

[1] Hamshire, C., Jack, K., Forsyth, R., Langan, A. M., & Harris, W. E. (2019). The wicked problem of healthcare student attrition. Nursing Inquiry, 26(3), e12294. doi:10.1111/nin.12294.

[2] Webber, D. A., & Ehrenberg, R. G. (2010). Do expenditures other than instructional expenditures affect graduation and persistence rates in American higher education? Economics of Education Review, 29(6), 947–958. doi:10.1016/j.econedurev.2010.04.006.

[3] Arias, A., Linares-Vásquez, M., & Héndez-Puerto, N. R. (2024). Undergraduate dropout in Colombia: A systematic literature review of causes and solutions. Journal of Latinos and Education, 23(2), 612–627. doi:10.1080/15348431.2023.2171042.

[4] Graham, M. J., Frederick, J., Byars-Winston, A., Hunter, A. B., & Handelsman, J. (2013). Increasing persistence of college students in STEM. Science, 341(6153), 1455–1456. doi:10.1126/science.1240487.

[5] Eltanahy, M., Forawi, S., & Mansour, N. (2020). Incorporating entrepreneurial practices into STEM education: Development of interdisciplinary E-STEM model in high school in the United Arab Emirates. Thinking Skills and Creativity, 37, 100697. doi:10.1016/j.tsc.2020.100697.

[6] Husain, F. Y., Forawi, S., & Chang, C. Y. (2023). Triple helix components supporting STEM education to increase future STEM careers in the United Arab Emirates. EURASIA Journal of Mathematics, Science and Technology Education, 19(8), em2303. doi:10.29333/ejmste/13424.

[7] Tlepina, S., Sarsembayev, M., Abaideldinov, Y., Balmagambetova, V., & Zukay, Z. (2024). A New Concept of Specialized Standards to Improve the Quality of Higher Legal Education. Emerging Science Journal, 8(4), 1385–1401. doi:10.28991/ESJ-2024-08-04-09.

[8] Muysken, J., & Nour, S. (2006). Deficiencies in education and poor prospects for economic growth in the Gulf countries: The case of the UAE. The Journal of Development Studies, 42(6), 957–980. doi:10.1080/00220380600774756.

[9] Ashour, S. (2020). Analysis of the attrition phenomenon through the lens of university dropouts in the United Arab Emirates. Journal of Applied Research in Higher Education, 12(2), 357–374. doi:10.1108/JARHE-05-2019-0110.

[10] Ifenthaler, D., & Yau, J. Y. K. (2020). Utilising learning analytics to support study success in higher education: a systematic review. Educational Technology Research and Development, 68(4), 1961–1990. doi:10.1007/s11423-020-09788-z.

[11] Marbouti, F., Diefes-Dux, H. A., & Madhavan, K. (2016). Models for early prediction of at-risk students in a course using standards-based grading. Computers & Education, 103, 1–15. doi:10.1016/j.compedu.2016.09.005.

[12] Salas-Pilco, S. Z., & Yang, Y. (2022). Artificial intelligence applications in Latin American higher education: A systematic review. International Journal of Educational Technology in Higher Education, 19(1), 21. doi:10.1186/s41239-022-00326-w.

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

[14] Luan, H., & Tsai, C. C. (2021). A review of using machine learning approaches for precision education. Educational Technology & Society, 24(1), 250–266.

[15] Riestra-González, M., Paule-Ruíz, M. P., & Ortin, F. (2021). Massive LMS log data analysis for the early prediction of course-agnostic student performance. Computers & Education, 163, 104108. doi:10.1016/j.compedu.2020.104108.

[16] Leal, F., Veloso, B., Pereira, C. S., Moreira, F., Durão, N., & Silva, N. J. (2022). Interpretable success prediction in higher education institutions using pedagogical surveys. Sustainability, 14(20), 13446. doi:10.3390/su142013446.

[17] Atif, A., Richards, D., Liu, D., & Bilgin, A. A. (2020). Perceived benefits and barriers of a prototype early alert system to detect engagement and support 'at-risk' students: The teacher perspective. Computers & Education, 156, 103954. doi:10.1016/j.compedu.2020.103954.

[18] Johora, F. T., Hasan, M. N., Rajbongshi, A., Ashrafuzzaman, M., & Akter, F. (2025). An explainable AI-based approach for predicting undergraduate students academic performance. Array, 26, 100384. doi:10.1016/j.array.2025.100384.

[19] Mustofa, S., Emon, Y. R., Mamun, S. B., Akhy, S. A., & Ahad, M. T. (2025). A novel AI-driven model for student dropout risk analysis with explainable AI insights. Computers and Education: Artificial Intelligence, 8, 100352. doi:10.1016/j.caeai.2024.100352.

[20] Zanellati, A., Zingaro, S. P., & Gabbrielli, M. (2024). Balancing performance and explainability in academic dropout prediction. IEEE Transactions on Learning Technologies, 17, 2086–2099. doi:10.1109/TLT.2024.3425959.

[21] Kiss, B., Nagy, M., Molontay, R., & Csabay, B. (2019). Predicting dropout using high school and first-semester academic achievement measures. In: 2019 17th International Conference on Emerging eLearning Technologies and Applications (ICETA); 2019 Nov 21; pp. 383–389. IEEE. doi:10.1109/ICETA48886.2019.9040158.

[22] Alsubhi, B., Alharbi, B., Aljojo, N., Banjar, A., Tashkandi, A., Alghoson, A., & Al-Tirawi, A. (2023). Effective feature prediction models for student performance. Engineering, Technology & Applied Science Research, 13(5), 11937–11944. doi:10.48084/etasr.6345.

[23] Albreiki, B., Habuza, T., & Zaki, N. (2022). Framework for automatically suggesting remedial actions to help students at risk based on explainable ML and rule-based models. International Journal of Educational Technology in Higher Education, 19(1), 49. doi:10.1186/s41239-022-00354-6.

[24] Oqaidi, K., Aouhassi, S., & Mansouri, K. (2022). Towards a students' dropout prediction model in higher education institutions using machine learning algorithms. International Journal of Emerging Technologies in Learning (iJET), 17(18), 103–117. doi:10.3991/ijet.v17i18.25567.

[25] Ebrahim, P., Al-Moumni, M., Al-Hattami, A., & Ali, A. (2021). A study of student attrition in the foundation year program of a teachers' college. International Journal of Lifelong Education, 40(3), 198–214. doi:10.1080/02601370.2021.1931973.

[26] Alteneiji, E. (2023). Value changes in gender roles: Perspectives from three generations of Emirati women. Cogent Social Sciences, 9(1), 2184899. doi:10.1080/23311886.2023.2184899.

[27] Ben Said, M., Hadj Kacem, Y., Algarni, A., & Masmoudi, A. (2024). Early prediction of student academic performance based on machine learning algorithms: A case study of bachelor's degree students in KSA. Education and Information Technologies, 29(11), 13247–13270. doi:10.1007/s10639-023-12370-8.

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

[29] Tinto, V. (1993). Leaving College: Rethinking the Causes and Cures of Student Attrition (2nd Edi.). University of Chicago Press, Chicago, United States.

[30] Al Murshidi, G. (2019). STEM education in the United Arab Emirates: Challenges and possibilities. International Journal of Learning, Teaching and Educational Research, 18(12), 316–332. doi:10.26803/ijlter.18.12.18.

[31] Yaghi, A., & Alabed, N. (2025). Factors affecting university dropout: Comparison of STEM and public affairs and management students. International Journal of Public Administration, 1–15. doi:10.1080/01900692.2025.2476676.

[32] Shafiq, D. A., Marjani, M., Habeeb, R. A. A., & Asirvatham, D. (2022). Student retention using educational data mining and predictive analytics: A systematic literature review. IEEE Access, 10, 72480–72503. doi:10.1109/ACCESS.2022.3188767.

[33] Aina, C., Baici, E., Casalone, G., & Pastore, F. (2022). The determinants of university dropout: A review of the socio-economic literature. Socio-Economic Planning Sciences, 79, 101102. doi:10.1016/j.seps.2021.101102.

[34] Lorenzo-Quiles, O., Galdón-López, S., & Lendínez-Turón, A. (2023). Factors contributing to university dropout: A review. Frontiers in Education, 8, 1159864. doi:10.3389/feduc.2023.1159864.

[35] Naseem, M., Chaudhary, K., & Sharma, B. (2022). Predicting freshmen attrition in computing science using data mining. Education and Information Technologies, 27(7), 9587–9617. doi:10.1007/s10639-022-11018-3.

[36] Demeter, E., Dorodchi, M., Al-Hossami, E., Benedict, A., Slattery Walker, L., & Smail, J. (2022). Predicting first-time-in-college students' degree completion outcomes. Higher Education, 1–21. doi:10.1007/s10734-021-00790-9.

[37] Bañeres, D., Rodríguez-González, M. E., Guerrero-Roldán, A. E., & Cortadas, P. (2023). An early warning system to identify and intervene online dropout learners. International Journal of Educational Technology in Higher Education, 20(1), 3. doi:10.1186/s41239-022-00371-5.

[38] Islam, M. M., Sojib, F. H., Mihad, M. F. H., Hasan, M., & Rahman, M. (2025). The integration of explainable AI in educational data mining for student academic performance prediction and support system. Telematics and Informatics Reports, 100203. doi:10.1016/j.teler.2025.100203.

[39] Abdulghani, H. M., Alanazi, K., Alotaibi, R., Alsubeeh, N. A., Ahmad, T., & Haque, S. (2023). Prevalence of potential dropout thoughts and their influential factors among Saudi medical students. Sage Open, 13(1), 21582440221146966. doi:10.1177/21582440221146966.

[40] Hammoudi Halat, D., Abdel-Salam, A. S. G., Bensaid, A., Soltani, A., Alsarraj, L., Dalli, R., & Malki, A. (2023). Use of machine learning to assess factors affecting progression, retention, and graduation in first-year health professions students in Qatar: a longitudinal study. BMC Medical Education, 23(1), 909. doi:10.1186/s12909-023-04887-w.

[41] Verbert, K., Ochoa, X., De Croon, R., Dourado, R. A., & De Laet, T. (2020). Learning analytics dashboards: The past, the present and the future. In: Proceedings of the Tenth International Conference on Learning Analytics & Knowledge; 2020 Mar; pp. 35–40. doi:10.1145/3375462.337550.

[42] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 16, 321–357. doi:10.1613/jair.953.

[43] Lunardon, N., Menardi, G., & Torelli, N. (2014). ROSE: A package for binary imbalanced learning. The R Journal, 6(1), 79–89. doi:10.32614/RJ-2014-008.

[44] Helland, H., & Hovdhaugen, E. (2022). Degree completion in short professional courses: Does family background matter? Journal of Further and Higher Education, 46(5), 680–694. doi:10.1080/0309877X.2021.1998394.

[45] Rizvi, S., Rienties, B., & Khoja, S. A. (2019). The role of demographics in online learning: A decision tree based approach. Computers & Education, 137, 32–47. doi:10.1016/j.compedu.2019.04.001.

[46] Nikolaidis, P., Ismail, M., Shuib, L., Khan, S., & Dhiman, G. (2022). Predicting student attrition in higher education through the determinants of learning progress: A structural equation modelling approach. Sustainability, 14(20), 13584. doi:10.3390/su142013584.

[47] Segura, M., Mello, J., & Hernández, A. (2022). Machine learning prediction of university student dropout: Does preference play a key role? Mathematics, 10(18), 3359. doi:10.3390/math10183359.

[48] Bernacki, M. L., Chavez, M. M., & Uesbeck, P. M. (2020). Predicting achievement and providing support before STEM majors begin to fail. Computers & Education, 158, 103999. doi:10.1016/j.compedu.2020.103999.

[49] Vidal, J., Gilar-Corbi, R., Pozo-Rico, T., Castejón, J. L., & Sánchez-Almeida, T. (2022). Predictors of university attrition: Looking for an equitable and sustainable higher education. Sustainability, 14(17), 10994. doi:10.3390/su141710994.

[50] Li, I. W., Jackson, D., & Koshy, P. (2024). Student's reported satisfaction at university: The role of personal characteristics and secondary school background. Higher Education, 1–19. doi:10.1007/s10734-024-01286-y.

[51] Alshamaila, Y., Alsawalqah, H., Aljarah, I., Habib, M., Faris, H., Alshraideh, M., & Salih, B. A. (2024). An automatic prediction of students' performance to support the university education system: a deep learning approach. Multimedia Tools and Applications, 83(15), 46369–46396. doi:10.1007/s11042-024-18262-4.

[52] Kocsis, Á., & Molnár, G. (2025). Factors influencing academic performance and dropout rates in higher education. Oxford Review of Education, 51(3), 414–432. doi:10.1080/03054985.2024.2316616.

[53] Lee, S., & Chung, J. Y. (2019). The machine learning-based dropout early warning system for improving the performance of dropout prediction. Applied Sciences, 9(15), 3093. doi:10.3390/app9153093.

[54] White, B. A., Miles, J. R., & Frantell, K. A. (2021). Intergroup dialogue: A justice‐ centered pedagogy to address gender inequity in STEM. Science Education, 105(5), 1010-1034. doi:10.1002/sce.21599.

[55] Vooren, M., Haelermans, C., Groot, W., & van den Brink, H. M. (2022). Comparing success of female students to their male counterparts in the STEM fields: An empirical analysis from enrollment until graduation using longitudinal register data. International Journal of STEM Education, 9(1), 1. doi:10.1186/s40594-021-00318-8.

[56] Montmarquette, C., Mahseredjian, S., & Houle, R. (2001). The determinants of university dropouts: a bivariate probability model with sample selection. Economics of education review, 20(5), 475–484. doi:10.1016/S0272-7757(00)00029-7.

[57] Deho, O. B., Joksimovic, S., Li, J., Zhan, C., Liu, J., & Liu, L. (2023). Should learning analytics models include sensitive attributes? Explaining the why. IEEE Transactions on Learning Technologies, 16(4), 560–572. doi:10.1109/TLT.2022.3226474.

[58] Azizah, Z., Ohyama, T., Zhao, X., Ohkawa, Y., & Mitsuishi, T. (2024). Predicting at-risk students in the early stage of a blended learning course via machine learning using limited data. Computers and Education: Artificial Intelligence, 7, 100261. doi:10.1016/j.caeai.2024.100261.