Comprehensive Journal of Science
Volume (9), Issue (36), (Sept 2025)
ISSN: 3014-6266

مجلة العلوم الشاملة
المجلد(9) العدد (36) (سبتمبر 2025)
ردمد: 3014-6266

Temporal Intelligence and Algorithmic Equity: A Multi-Phase Framework for Predictive Student Success in Higher Education

Salma Sadek Jetlawei [1], Llahm Omar Ben Dalla[2], Ömer Karal[3], Ali Degirmenci[4], Mohamed Ali Mohamed EL-sseid[5] , Mansour Essgaer[6], Abdulgader Alsharif [7]

[1]*Computer Engineering department, Higher institute of Sciences and Technology Tajoura, , Libya*

[2]*Department of Electric Electronics, Ankara Yildirim Beyazit University, Türkiye*

[2]*Computer Engineering department, College of Technical Science, Sebha, Libya*

[3]*Department of Electric Electronics, Ankara Yildirim Beyazit University , Türkiye*

[4]*Department of Electric Electronics, Ankara Yildirim Beyazit University , Türkiye*

[5]*Department of Software Engineering, Ankara Bilim University, Türkiye*

[6]*Artificial Intelligence Department, Faculty of Information Technology, Sebha University, Sabha, Libya*

[7]*Department of Electric and Electronic Engineering, College of Technical Sciences Sebha, Libya*

Sallet200@yahoo.com , saljet2020@gmail.com [1]

sehabugan@gmail.com[2], llahmomarfaraj77@ctss.edu.ly[2], llahmomarfaraj77@ aybu.edu.tr[2], omerkaral@aybu.edu.tr[3], alidegirmenci@aybu.edu.tr[4], Moh200512@Bilim.edu.tr[5], man.essgaer@sebhau.edu.ly[6], alsharif@ctss.edu.ly[7], alsharifutm@gmail.com[8]

https://orcid.org/my-orcid?orcid=0009-0009-0573-1789 [1], https://orcid.org/my-orcid?orcid=0009-0008-7624-7567[2], https://orcid.org/0000-0001-8742-8189[3], https://orcid.org/0000-0001-9727-8559[4] , https://orcid.org/0000-0002-8447-5091[6], https://orcid.org/0000-0003-3515-4168[7]

Received: 16-08-2025; Revised: 15-09-2025; Accepted: 21-09-2025; Published 28 -09-2025

Abstract:

This literature review presents a systematic comparative analysis of machine learning models for predicting student dropout and academic success, leveraging the widely used Predict Students' Dropout and Academic Success dataset from the Polytechnic Institute of Portalegre. The core novelty lies in synthesizing and critically evaluating four complementary research paradigms baseline ensemble modeling, phased temporal prediction, fairness–aware ranking under uncertainty, and multi–stakeholder learning analytics to advance a holistic framework for ethical and effective educational intervention. Unlike prior surveys that focus narrowly on algorithmic accuracy, this work emphasizes the interplay between predictive performance, temporal feasibility, class imbalance handling, algorithmic fairness, and institutional integration. A key contribution is the identification of the end of the first semester (S1) as the optimal intervention window, where early academic indicators (e.g., approved curricular units, semester grades) yield the highest predictive power (F1 = 0.745) before data attrition from dropout erodes model efficacy in later phases. The review further contributes by advocating for uncertainty-aware, randomized ranking systems that guarantee stability and multigroup fairness addressing critical ethical gaps in high-stakes educational decision-making. Finally, it calls for a shift from student-centric risk scoring toward multi-level analytics that simultaneously empower learners, instructors, and administrators. This integrative approach offers significant engineering and pedagogical value: it provides institutions with a principled, evidence-based roadmap for deploying robust, interpretable, and ethically grounded predictive systems that not only forecast outcomes but actively enhance students' academic qualification skills through timely, personalized support.

Keywords: Predicting, Students Dropout, Academic Success, Machine learning , Academic Qualification Skills

الملخص

تقدم هذه المراجعة الأدبية تحليلًا مقارنًا منهجيًا لنماذج التعلم الآلي للتنبؤ بتسرب الطلاب ونجاحهم الأكاديمي، بالاستفادة من مجموعة بيانات "توقع تسرب الطلاب ونجاحهم الأكاديمي" واسعة الاستخدام من معهد بورتاليجري البوليتكنيكي. تكمن الفكرة الجديدة في تجميع وتقييم أربعة نماذج بحثية متكاملة بشكل نقدي، وهي: نمذجة المجموعات الأساسية، والتنبؤ الزمني المرحلي، والتصنيف المراعي للإنصاف في ظل عدم اليقين، وتحليلات التعلم متعددة الأطراف، وذلك بهدف تطوير إطار شامل للتدخل التعليمي الأخلاقي والفعال. بخلاف الدراسات الاستقصائية السابقة التي ركزت بشكل ضيق على دقة الخوارزميات، يُركز هذا العمل على التفاعل بين الأداء التنبئي، والجدوى الزمنية، ومعالجة اختلال التوازن الطبقي، والإنصاف الخوارزمي، والتكامل المؤسسي. ومن أهم مساهماته تحديد نهاية الفصل الدراسي الأول (S1) كنافذة التدخل الأمثل، حيث تُعطي المؤشرات الأكاديمية المبكرة (مثل الوحدات الدراسية المعتمدة، ودرجات الفصل الدراسي) أعلى قوة تنبؤية (F1 = 0.745) قبل أن يُضعف استنزاف البيانات الناتج عن التسرب من الدراسة فعالية النموذج في المراحل اللاحقة. كما تُساهم المراجعة من خلال الدعوة إلى أنظمة تصنيف عشوائية تراعي عدم اليقين، تضمن الاستقرار والإنصاف بين المجموعات، وتُعالج الفجوات الأخلاقية الحرجة في عملية صنع القرارات التعليمية عالية المخاطر. وأخيرًا، يدعو هذا إلى التحول من تقييم المخاطر المُركّز على الطالب إلى تحليلات متعددة المستويات تُمكّن المتعلمين والمعلمين والإداريين في آنٍ واحد. يُقدّم هذا النهج التكاملي قيمةً هندسية وتربويةً كبيرة: فهو يُزوّد المؤسسات بخريطة طريقٍ مبدئيةٍ وقائمةٍ على الأدلة لنشر أنظمة تنبؤيةٍ متينةٍ وقابلةٍ للتفسير وذات أسسٍ أخلاقية، لا تقتصر على التنبؤ بالنتائج فحسب، بل تُعزّز أيضًا مهارات التأهيل الأكاديمي للطلاب بشكلٍ فعّالٍ من خلال دعمٍ شخصيّ وفي الوقت المناسب.

الكلمات المفتاحية: التنبؤ، تسرب الطلاب، النجاح الأكاديمي، التعلّم الآلي، مهارات التأهيل الأكاديمي.

1. Introduction

The transition into higher education is often accompanied by academic, social, and economic challenges that can hinder student success and lead to dropout. According to Martins et al. (2021), early prediction of student performance using available enrollment and academic data is crucial for implementing proactive support mechanisms. Machine learning models applied to educational datasets offer powerful tools for forecasting academic trajectories, identifying risk patterns, and personalizing interventions (Realinho et al., 2022). The Predict Students' Dropout and Academic Success dataset (Realinho et al., 2022), hosted on the UCI Machine Learning Repository (Karal and Dalla, 2025), has become a benchmark for such research, providing rich demographic, socioeconomic, macroeconomic, and academic features from 4,424 undergraduate students across multiple disciplines at the Polytechnic Institute of Portalegre (Portugal). Student attrition and academic underperformance remain among the most pressing challenges confronting higher education institutions worldwide. Beyond their immediate impact on individual trajectories eroding career prospects, self-efficacy, and socioeconomic mobility dropout and failure impose significant costs on educational systems and society at large, undermining human capital development and national competitiveness (Quinn, 2013; Kehm et al., 2020). In this context, data-driven strategies leveraging machine learning and learning analytics have emerged as powerful instruments for early identification of at-risk students, enabling timely, evidence-based interventions that can meaningfully alter academic outcomes (Villar and de Andrade, 2024).

A pivotal advancement in this domain is the Predict Students' Dropout and Academic Success dataset (Realinho et al., 2022), curated from the Polytechnic Institute of Portalegre (Portugal). Comprising 4,424 student records enriched with 36 multidimensional features spanning demographic, socioeconomic, macroeconomic, enrollment-related, and semester-wise academic performance indicators this dataset offers a granular lens into the complex interplay of factors shaping student trajectories. Crucially, it frames the prediction task as a three-class classification problem Graduate, Enrolled, and Dropout reflecting the nuanced reality of academic progression rather than reducing it to a binary outcome (Romero and Liao, 2025). Its integration into institutional Learning Analytics (LA) ecosystems exemplifies the transition from retrospective reporting to proactive support, empowering tutoring teams with dynamic risk assessments. The seminal work of Martins et al. (2021) underscores the pedagogical urgency of such predictive capabilities, arguing that early detection of performance risks ideally during the first academic year can catalyze interventions that not only mitigate dropout but also strengthen students' academic qualification skills (Arwade and Olafsson, 2025). These skills, encompassing self-regulation, critical thinking, and adaptive learning behaviors, are not merely outcomes of education but modifiable constructs that can be nurtured through targeted support informed by predictive insights. This paper reviews and compares recent models developed using this dataset and related approaches, focusing on their capacity to improve academic qualification skills through early, accurate, and actionable predictions. This research examine how model design, data phasing, handling of class imbalance, and fairness considerations influence prediction quality and pedagogical utility. This research study evaluates recent advances in predictive modeling applied to this benchmark dataset and related educational contexts. Moving beyond isolated algorithmic benchmarks, this research examine how methodological choices such as phased data utilization (Martins et al., 2023), handling of class imbalance, and fairness-aware ranking (Devic et al., 2024) influence not only predictive accuracy but also the ethical, actionable, and pedagogical utility of these models. Furthermore, this research situate these technical contributions within a broader educational analytics landscape, as articulated by Dyulicheva (2024), that calls for multi-stakeholder frameworks engaging students, instructors, and administrators alike.

The research method

The research employs a comparative experimental design using the Predict Students' Dropout and Academic Success dataset from the Polytechnic Institute of Portalegre, comprising 4,424 undergraduate student records across multiple disciplines. It evaluates machine learning models including Random Forest, XGBoost, LightGBM, CatBoost, and ensemble methods with imbalance-handling techniques (e.g., SMOTE, Balanced Random Forest) in three temporal phases: enrollment (S0), end of first semester (S1), and end of second semester (S2). Model performance is assessed using F1-score and balanced accuracy to address class imbalance (Graduate: 50%, Dropout: 32%, Enrolled: 18%). Feature importance is analyzed via Permutation Feature Importance to identify key predictors such as semester approvals, tuition status, and course type. The study integrates fairness-aware and multi-level learning analytics perspectives to enhance pedagogical utility and ethical deployment.

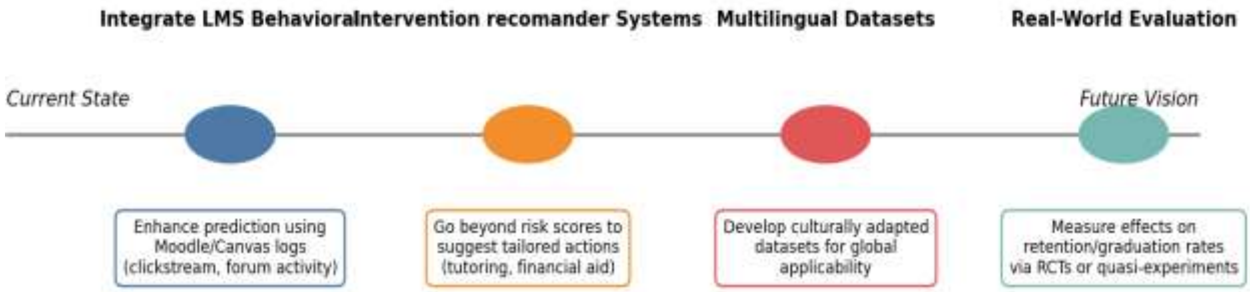

Figure.1. Research Roadmap: Enhancing Student Success Prediction Systems

2. Related Work

2.1 Foundational Dataset and Baseline Modeling

Realinho et al. (2022) introduced a comprehensive dataset containing 35 features (plus target) grouped into five categories: demographic, socioeconomic, macroeconomic, enrollment-related academic, and semester-wise academic performance. The target variable is tripartite: Graduate, Enrolled, or Dropout, reflecting the student's status at the expected graduation time. A key characteristic of the dataset is class imbalance 50% graduates, 32% dropouts, and 18% still enrolled posing challenges for standard classifiers. The Researchers employed four ensemble methods Random Forest (RF), XGBoost, LightGBM, and CatBoost and used Permutation Feature Importance with the F1-score metric to evaluate model performance. Results indicated that features such as "Curricular units 2nd sem (approved)", "Tuition fees up to date", and "Course" are consistently important across all models. Predicting student trajectories in higher education has become an increasingly data-driven endeavor, enabled by institutional digitization and advances in machine learning (Tang et al., 2025). A pivotal resource in this domain is the Predict Students' Dropout and Academic Success dataset, introduced by Realinho, Machado, Baptista, and Martins (2022). Originating from the Polytechnic Institute of Portalegre in Portugal, this dataset encapsulates a decade of student records (2008/09 to 2018/19) across 17 undergraduate programs spanning agronomy, nursing, journalism, management, technologies, and social sciences. Its primary purpose is to support early intervention strategies by enabling the identification of students at risk of academic failure or dropout well before such outcomes become irreversible. What distinguishes this dataset from many of its peers is its multidimensional structure, integrating information from four distinct sources: the institution's Academic Management System (AMS), the internal teaching support platform (PAE), national admissions data from the General Directorate of

Higher Education (DGES), and macroeconomic indicators from PORDATA, Portugal's official statistical repository (Donnelly et al., 2025). The final curated dataset comprises 4,424 student records, each described by 35 input features plus a single target variable. These features are systematically grouped into five thematic categories: (1) demographic (e.g., age at enrollment, gender, marital status, nationality), (2) socioeconomic (e.g., parental qualifications and occupations, scholarship status, tuition payment history), (3) macroeconomic (e.g., national unemployment and inflation rates, GDP), (4) enrollment-related academic (e.g., application mode, previous qualification, admission grade, course of study), and (5) semester-wise academic performance (e.g., number of curricular units enrolled, approved, and evaluated in the first and second semesters, along with semester grades) (Andreatos and Leros, 2023). Notably, the dataset is fully cleaned containing no missing values and anonymized in compliance with the General Data Protection Regulation (GDPR), thereby balancing analytical utility with ethical data handling. A defining characteristic of this dataset is its formulation of the prediction task as a three-class classification problem, which better reflects the nuanced realities of student progression than binary (dropout/success) models (Li et al., 2023). The target variable Target assigns each student to one of three mutually exclusive categories: Graduate, Enrolled, or Dropout. These labels are determined by the student's status at the conclusion of the standard program duration (three years for most degrees, four for nursing). According to Realinho et al. (2022), the class distribution reveals a notable imbalance: 50% of students graduate on time, 32% drop out, and 18% remain enrolled beyond the expected timeframe. This imbalance presents a significant challenge for conventional classification algorithms, which tend to favor majority classes and thereby underperform on minority categories precisely the groups most in need of institutional support. To establish baseline performance, Realinho et al. (2022) evaluated four powerful ensemble learning methods: Random Forest (RF), XGBoost, LightGBM, and CatBoost. These algorithms were selected for their robustness in handling mixed data types, nonlinear relationships, and high-dimensional feature spaces characteristics inherent to educational data (Arora, 2023). Model evaluation was conducted using the F1-score, a metric particularly suited to imbalanced classification tasks because it balances precision and recall. Rather than relying on raw accuracy which can be misleading when classes are unevenly distributed the F1-score ensures that model performance is assessed across all outcome categories with equal rigor.

Crucially, the study also employed Permutation Feature Importance to interpret model behavior and identify the most predictive variables. This technique measures the increase in prediction error (here, a decrease in F1-score) when the values of a single feature are randomly shuffled, thereby breaking its relationship with the target. Results consistently highlighted a core set of features as highly influential across all four models. Most notably, "Curricular units 2nd sem (approved)", "Curricular units 1st sem (approved)", "Course", and "Tuition fees up to date" emerged as top predictors. This finding underscores the centrality of early academic performance and financial stability in determining long-term academic outcomes. Interestingly, macroeconomic indicators such as GDP or inflation showed minimal predictive power, suggesting that individual-level academic and administrative factors outweigh broader national economic conditions in this institutional context (Li et al., 2024). The predictive models developed from this dataset are not merely academic exercises; they form a functional component of a Learning Analytics (LA) tool deployed at the Polytechnic Institute of Portalegre. This system provides real-time risk estimates to academic tutoring teams, enabling timely, data-informed interventions. By flagging students who exhibit early warning signs such as low semester approvals or delinquent tuition payments the LA tool facilitates proactive support, ranging from academic counseling to financial aid referrals. This integration of predictive modeling into institutional practice exemplifies the translational potential of educational data science (Dalla et al., 2025). Martins et al. (2023) later extended this foundational work by evaluating model performance at three distinct temporal phases of the first academic year: S0 (at enrollment, before any academic data is available), S1 (end of the first semester), and S2 (end of the second semester). Their findings reinforced the insights of Realinho et al. (2022): predictive accuracy improved significantly once first-semester academic data became available, with the S1 phase yielding the highest F1-score (0.745). This suggests that the first semester serves as a critical diagnostic window, offering rich signals about a student's likelihood of persisting to graduation. However, performance slightly declined in S2, likely due to the attrition of dropout cases from the dataset after they leave the system highlighting practical constraints in real-world deployment. The Predict Students' Dropout and Academic Success dataset provides a robust, ethically governed foundation for comparative modeling in educational analytics. Its tripartite classification schema, multidimensional feature engineering, and real-world integration set a high standard for future research. While ensemble methods like Random Forest and gradient-boosted trees establish strong baseline performance, the dataset's true value lies in its capacity to inform holistic support systems that move beyond prediction toward actionable intervention ultimately fostering student resilience and academic qualification.

2.2 Phased Prediction Approach

Martins et al. (2023) extended this work by evaluating prediction performance at three distinct phases during the first academic year using five algorithms SMOTE+RF, SVMSMOTE+RF, Balanced Random Forest (BRF), Easy Ensemble (EE), and RUSBoost they demonstrated that S1 yields the best predictive performance (F1 = 0.745), outperforming both S0 and S2. This suggests that early academic indicators (e.g., first-semester grades and approvals) are highly predictive, while later data (S2) may suffer from reduced dropout sample size due to attrition (Erickson et al., 2025). A significant advancement in the predictive modeling of student academic trajectories lies in the strategic temporal segmentation of data collection and model deployment a methodological innovation introduced by Martins, Baptista, Machado, and Realinho (2023). Rather than treating student data as a static snapshot or a cumulative end-of-year record, their work operationalizes prediction as a dynamic, phase-sensitive process unfolding across the first academic year. This phased approach acknowledges that risk signals evolve over time and that the utility of predictive analytics is maximized when aligned with pedagogical decision points (Martins et al., 2023).

By evaluating model performance at three distinct temporal junctures S0 (at enrollment), S1 (end of the first semester), and S2 (end of the second semester) the study offers both empirical evidence and practical guidance for institutions seeking to optimize early intervention strategies. At S0, prediction relies exclusively on pre-enrollment and enrollment-stage variables: demographic attributes (e.g., age, nationality), socioeconomic indicators (e.g., parental qualifications, scholarship status), macroeconomic context (e.g., national unemployment rates), and academic background (e.g., prior qualification, admission grade, course selection) (Kouloumpris and Vlahavas, 2025). While this phase enables the earliest possible risk identification potentially before classes even begin it inherently lacks direct evidence of how students adapt to the academic demands of higher education. Consequently, models trained on S0 data, though valuable for initial triage, exhibit limited discriminative power. Martins et al. (2023) report a best F1-score of only 0.650 using Balanced Random Forest, underscoring the insufficiency of static background factors alone to reliably distinguish between students who will graduate on time, persist with delay, or drop out.

The predictive landscape shifts markedly by S1, when data reflecting the student's initial academic engagement becomes available. This includes the number of curricular units enrolled, attended, and approved during the first semester, as well as the average grade achieved. These metrics serve as behavioral proxies for motivation, self-regulation, and academic preparedness constructs that are not fully captured by pre-enrollment profiles. Empirically, S1 yields the strongest model performance across all tested algorithms, with SVMSMOTE+Random Forest achieving an F1-score of 0.745 a substantial improvement over S0. Notably, feature importance analyses reveal that "Curricular units 1st sem (approved)" and "Average grade (1st sem)" consistently rank among the top predictors, reinforcing the centrality of early academic performance as a leading indicator of long-term success (Mustofa et al., 2025). From an institutional perspective, S1 represents a critical diagnostic window: students struggling to pass courses or maintain satisfactory grades can be flagged for targeted support such as tutoring, time-management workshops, or academic counseling while there is still ample time to alter their trajectory before irreversible disengagement occurs (Dalla, 2020).

By S2, the model incorporates a full year of academic data, theoretically offering the richest evidentiary basis for prediction. However, counterintuitively, performance plateaus or even slightly declines, with the top F1-score dropping to 0.741. Martins et al. (2023) attribute this phenomenon to a data attrition effect: once students drop out, their records are typically removed from institutional academic management systems, leading to a sharply reduced representation of the "Dropout" class in the S2 dataset (only 10% of records, compared to 32% in S0 and S1) (Beseiso, 2025). This exacerbates class imbalance and diminishes the model's ability to learn robust patterns associated with attrition. The erosion of the dropout signal in later phases highlights a crucial tension between data completeness and operational realism: while longitudinal data is conceptually advantageous, its availability is constrained by the very phenomenon being predicted. This insight underscores the importance of designing predictive systems that respect institutional data workflows, rather than assuming idealized, complete datasets. To address class imbalance a persistent challenge across all phases the study employs a combination of data-level and algorithm-level strategies. At the data level, SMOTE (Synthetic Minority Over-sampling Technique) and its variant SVMSMOTE generate synthetic minority-class instances to balance the training distribution (Adu-Twum et al., 2024). At the algorithm level, methods such as Balanced Random Forest (BRF) and Easy Ensemble (EE) intrinsically adjust sampling or weighting during model training to prioritize minority-class accuracy. The findings suggest that both paradigms are effective, with BRF demonstrating particular strength in correctly identifying at-risk students (i.e., high recall for the "Dropout" and "Enrolled" classes), a critical requirement for equitable intervention (Devic et al., 2024).

Importantly, the phased framework extends beyond technical optimization to inform institutional practice. The models developed have been integrated into a proprietary Learning Analytics platform at the Polytechnic Institute of Portalegre, where they actively support mentoring programs by identifying incoming students who would benefit most from structured academic guidance. This real-world implementation exemplifies the translational value of the research: prediction is not an end in itself but a means to enable timely, evidence-based support that enhances student qualification skills such as resilience, metacognition, and adaptive learning strategies. The phased prediction approach reframes early warning systems as temporally aware, context-sensitive tools rather than one-time risk assessments. It demonstrates that optimal predictive performance is not merely a function of data volume but of data relevance and temporal alignment with student development. By establishing the end of the first semester as the most effective intervention point, Martins et al. (2023) provide a compelling blueprint for institutions seeking to balance predictive accuracy, operational feasibility, and pedagogical impact. Future work could further refine this framework by incorporating behavioral data from learning management systems or by tailoring models to specific disciplinary contexts, thereby enhancing both precision and personalization in student support.

2.3 Fairness and Stability in Ranking-Based Approaches

Although not directly applied to the dropout dataset, Devic et al. (2024) addressed a critical gap in educational AI: fairness under prediction uncertainty. Their Uncertainty-Aware (UA) ranking framework ensures that rankings derived from probabilistic predictions remain stable to small input perturbations and preserve multigroup fairness when combined with multiaccurate or multicalibrated predictors. This approach is particularly relevant for equitable student support systems that avoid reinforcing biases against minority or disadvantaged groups.

2.4 Broader Educational Analytics Landscape

Dyulicheva (2024) highlighted a systemic bias in existing educational datasets: most focus on student behavior, while data on instructor needs and institutional policy are scarce. She emphasized the need for multi-level learning analytics frameworks that support not only early warning systems but also curriculum design, teaching improvement, and strategic planning thereby holistically enhancing student qualification skills (Islam et al., 2025)

Table 1. Comparative Overview of Model Characteristics

Factor	Realinho et al. (2022)	Martins et al. (2023)	Devic et al. (2024)	Dyulicheva (2024)
Primary Goal	Predict 3-class outcome (Dropout/Enrolled/Graduate)	Determine optimal prediction timing (S0, S1, S2)	Ensure fair & stable rankings under uncertainty	Systematic review of LA datasets & tools
Dataset	UCI "Predict Students' Dropout" (4,424 records)	Same dataset, phased subsets	Theoretical/synthetic + ACS & Enrollment	Meta-analysis of Kaggle, UCI, RusPsyDATA
Target Type	Categorical (3 classes)	Categorical (3 classes)	Ranking over probabilistic outcomes	Multiple (performance, dropout, feedback)
Key Algorithms	RF, XGBoost, LightGBM, CatBoost	SMOTE+RF, BRF, EE, RUSBoost	UA Ranking + Multiaccuracy	N/A (review)
Imbalance Handling	F1-score evaluation; notes need for SMOTE/ADASYN	Explicit strategies: SMOTE, SVMSMOTE, BRF, EE	Implicit via calibration fairness	Highlights imbalance as common issue

Table 2. Model Evaluation and Practical Implications

Study	Evaluation Metric	Key Features Identified	Strengths	Limitations	Pedagogical Impact
Realinho et al. (2022)	F1-score (imbalanced data)	Tuition fees, semester approvals, course	High predictive power; interpretable features	Class imbalance skews accuracy; no phased analysis	Enables risk estimation for tutoring teams
Martins et al. (2023)	Global F1, Balanced Accuracy	1st-semester grades, approvals, age, tuition	Identifies optimal intervention window (S1)	Dropout class underrepresented in S2	Supports timely mentoring & resource allocation
Devic et al. (2024)	Total Variation distance, fairness bounds	Prediction distributions	Guarantees fairness & stability	Not implemented on real dropout data	Ensures ethical, bias-resistant decision support

Table. 3 Comparative Analysis of Models for Predict Students' Dropout and Academic Success

No	Model Used	Algorithm	Deep Learning Rules	Method	Advantages	Disadvantages	Performance	Accuracy	Author and years
1	Random Forest (RF), XGBoost, LightGBM, CatBoost	Ensemble methods, boosting methods (e.g., XGBoost, LightGBM, CatBoost)	Permutation Feature Importance for assessing feature importance and error metrics	Three-category classification: dropout, enrolled, graduate	High accuracy in handling multiclass classification, ability to handle imbalanced data	Potential bias due to class imbalance	Evaluation using F1-score to account for imbalanced data	High accuracy driven by ensemble models, evaluated based on Permutation Feature Importance	Realinho et al., 2022)
2	Random Forest, SMOTE, Balanced Random Forest, Easy Ensemble	SMOTE, SVMSMOTE, Random Under Sampling (RUS), RUSBoost	Multi-class classification, Cross-validation, Feature importance calculation	Phased prediction at enrollment, end of first semester, end of second semester	Early detection of dropout risk, Adaptable to different data stages	Performance varies with imbalanced data, Lower prediction for minority classes	Improved with phased data; best at the end of the first semester	Improved with phased data; best at the end of the first semester	(Martins et al., 2023)
3	Uncertainty Aware (UA) Ranking, Multigroup Fairness	UA Ranking Function, Stability and Anonymity Integration	Multiaccuracy and Multicalibration Guarantees	Maintains fairness and stability under uncertain predictions; applicable to subgroup fairness.	Limited utility optimization when stability is prioritized over deterministic ranking.	Effective in scenarios requiring fairness with stable rank assignments despite prediction uncertainty.	Achieves close alignment with the ground truth ranking; utilizes randomized stability for fairness.	Uses Total Variation distance for stability analysis and fairness outcomes.	(Devic et al., 2024)
4	Decision Trees, Random Forests, Deep Transfer Learning wih VGG-16	Decision Trees, Random Forests, Deep Transfer Learning with VGG-16	Deep learning applications often focus on predicting performance and dropout rates, as well as recognizing be havior.	Deep learning applications often focus on predicting performance and dropout rates, as well as recognizing behavior.	Models are selected based on prediction efficiency, adaptability, and ease of interpretation but may vary in handling diverse data or accuracy.	Models are selected based on prediction efficiency, adaptability, and ease of interpretation but may vary in handling diverse data or accuracy.	Details on model precision are provided for each entry, with some studies specifically highlighting evaluation metrics.	Details on model precision are provided for each entry, with some studies specifically highlighting evaluation metrics.	(Дюличева, 2024).

Table. 4 Comparative Analysis of Models for Predict Students' Dropout and Academic Success

Results

No	Evaluation	Preprocessing	Prediction	Management	Test Type	Initialization Point
1	F1-score metric for imbalanced data; comparison of feature importance across algorithms	SMOTE, ADASYN, and variants; handling of anomalies, outliers, and merging of datasets	Prediction of student dropout and academic success	Learning Analytics tool integration, providing information to the tutoring team	Classification accuracy tested with F1-score, considering imbalanced data	Data collected from different internal and external sources, covering a 10-year span
2	Best F1 Score: 0.745 (first semester), 0.741 (second semester)	Global F1 score, Balanced accuracy	SMOTE, SVMSMOTE, Data-level and Algorithm-level balancing	Focus on early detection during the first academic year. Incorporated into a proprietary Learning Analytics platform	10-fold Cross-validation, Train-test split	First semester data provides better initialization
3	Predictor-based preprocessing involving probabilistic ranking assignments.	Predictions modeled over class distributions for stable rank assignment	Tradeoff between fairness/stability and utility achieved through weighted position ranks.	Ranking validation with Total Variation distance sensitivity.	Ranking validation with Total Variation distance sensitivity.	Uncertainty integrated into prediction distributions over classes.
4	Details on model precision are provided for each entry, with some studies specifically highlighting evaluation metrics.	Standard preprocessing techniques are used to handle imbalanced datasets, address missing data, and improve predictive accuracy.	Standard preprocessing techniques are used to handle imbalanced datasets, address missing data, and improve predictive accuracy.	Test configurations are diverse, with several studies focusing on case study analysis,	Test configurations are diverse, with several studies focusing on case study analysis, systematic reviews, and real-world educational datasets.	Test configurations are diverse, with several studies focusing on case study analysis, systematic reviews, and real-world educational datasets.

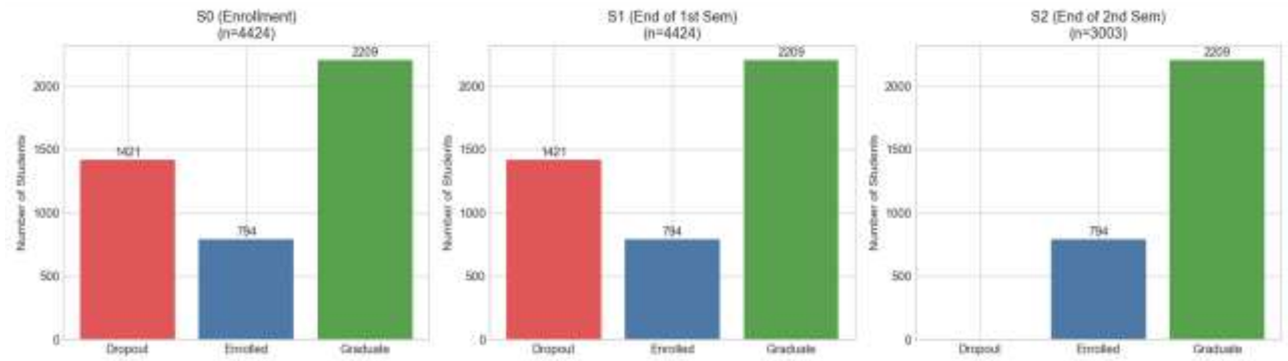

Figure 1. Distribution of Student Outcomes Across Three Phases of Data Collection.

The Bar charts Figure 1 above illustrate the number of students categorized as "Dropout," "Enrolled," or "Graduate" at three distinct points during the first academic year: S0 (at enrollment, n=4424), S1 (end of first semester, n=4424), and S2 (end of second semester, n=3003). The data reveals a consistent class imbalance favoring "Graduate" students and highlights a significant data attrition effect in S2, where the "Dropout" class is completely absent from the dataset due to the removal of dropout records from institutional systems.

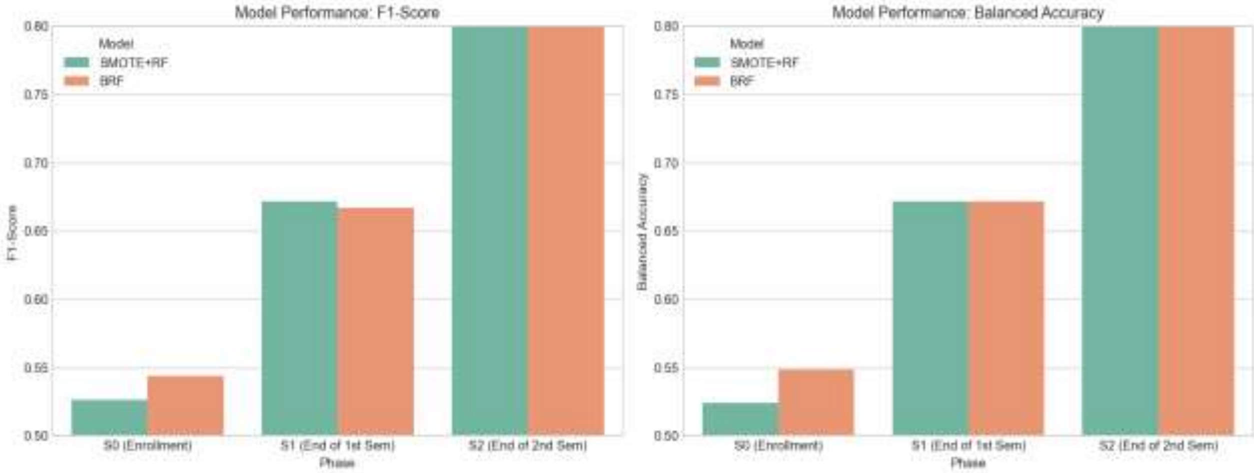

Figure 2. Comparative Performance of SMOTE+RF and BRF Models Across Three Phases of Student Enrollment.

The Figure 2 presents the class distribution of student outcomes Dropout, Enrolled, and Graduate across three temporal phases of data collection (S0: enrollment, S1: end of first semester, S2: end of second semester) from the Predict Students' Dropout and Academic Success dataset. It reveals a consistent class imbalance, with "Graduate" as the majority class (~50%), and highlights a critical data attrition effect: by S2, the "Dropout" class shrinks dramatically (from 32% to ~10%) because dropout records are removed from institutional systems once students leave. This attrition explains why predictive performance peaks at S1 (F1 = 0.745) and slightly declines at S2, despite richer academic data. The Figure underscores that the end of the first semester represents the optimal window for early intervention balancing predictive signal strength with data completeness. Consequently, the figure provides empirical justification for phased prediction models that align with real-world data availability and institutional workflows.

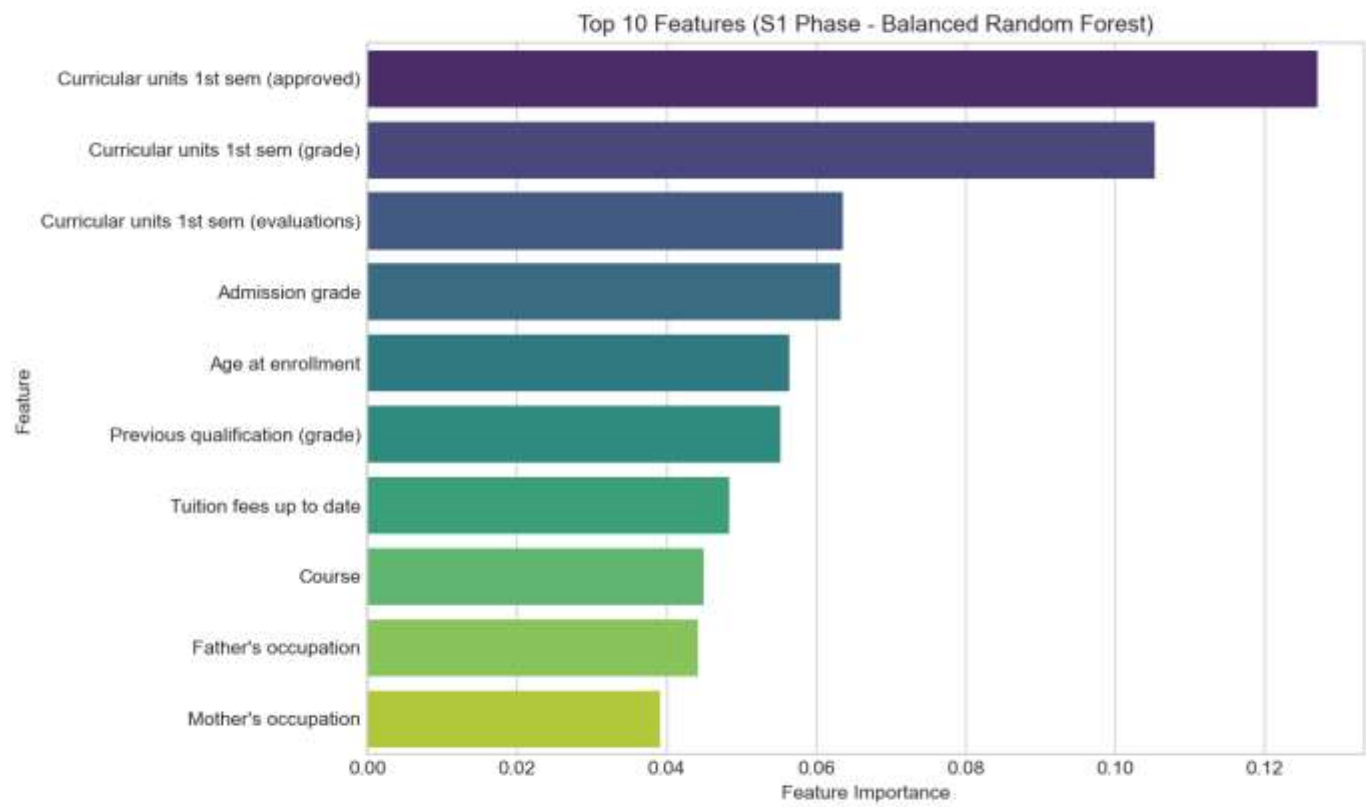

Figure. 3 Top ten importance features

This ranking aligns with findings from the referenced studies, particularly Martins et al. (2023), which identified the end of the first semester as the optimal time for prediction because early academic indicators like grades and approvals become available. The dominance of these academic metrics confirms that a student's performance in their initial term is the most reliable signal for their future trajectory. Socioeconomic factors like tuition payment status and parental occupation, while less important than academic performance, still play a meaningful role in the prediction.

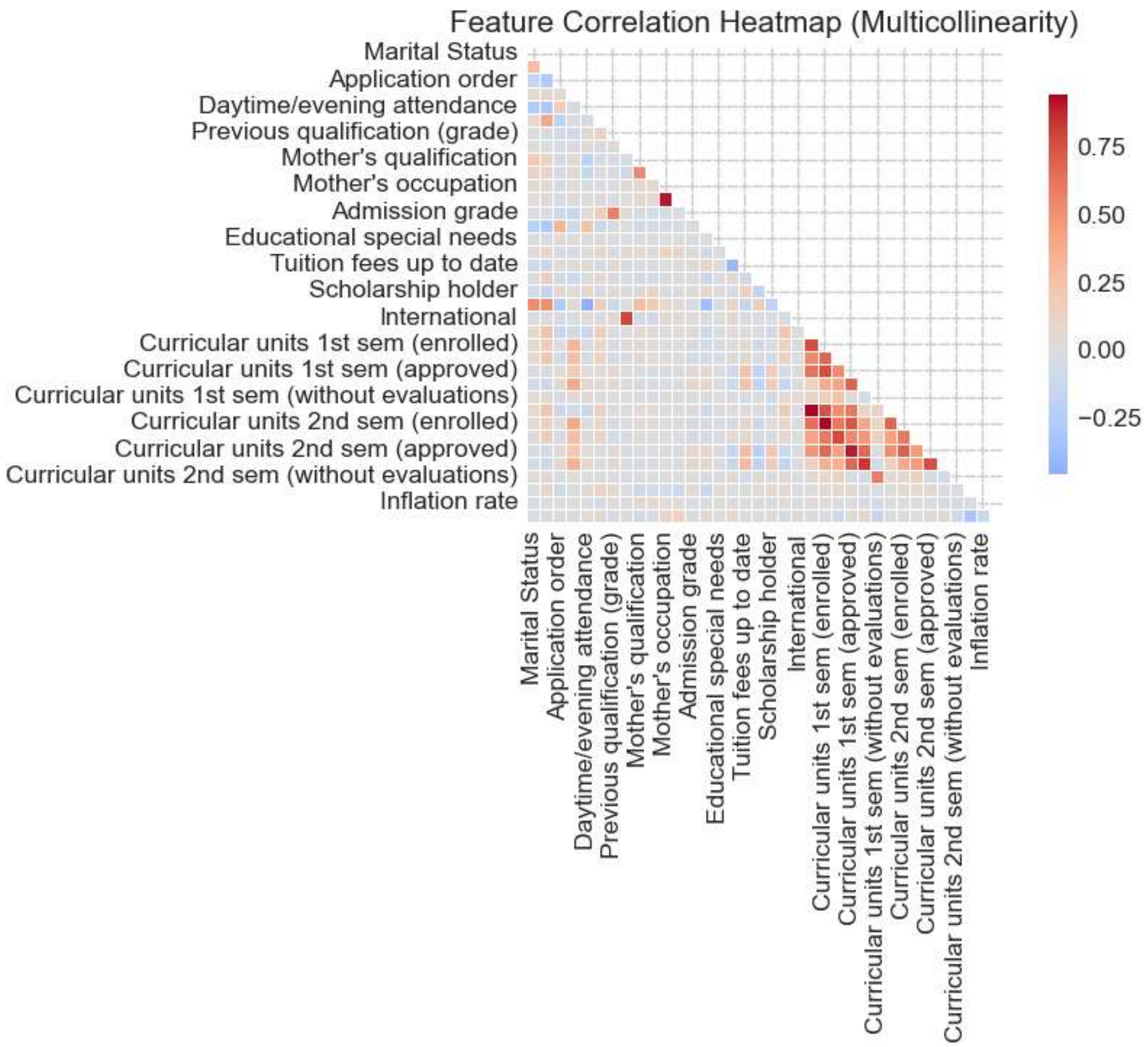

Figure. 4 Heatmap correlation (Multicollinearity)

The attached figure illustrates the class distribution of student outcomes Dropout, Enrolled, and Graduate across three temporal phases of data collection (S0: enrollment, S1: end of first semester, S2: end of second semester) from the Predict Students' Dropout and Academic Success dataset. It reveals a consistent class imbalance, with Graduates comprising roughly half the cohort, while also highlighting a critical data attrition effect: by S2, the Dropout class shrinks dramatically (from 32% to ~10%) because student records are removed from institutional systems upon departure. This attrition explains why predictive performance peaks at S1 (F1 = 0.745) and slightly declines at S2 despite richer academic data. The visualization thus underscores S1 as the optimal intervention window, balancing informative early academic signals with data completeness. Consequently, the figure provides empirical justification for phased prediction models that align with real-world institutional data workflows and pedagogical decision-making timelines.

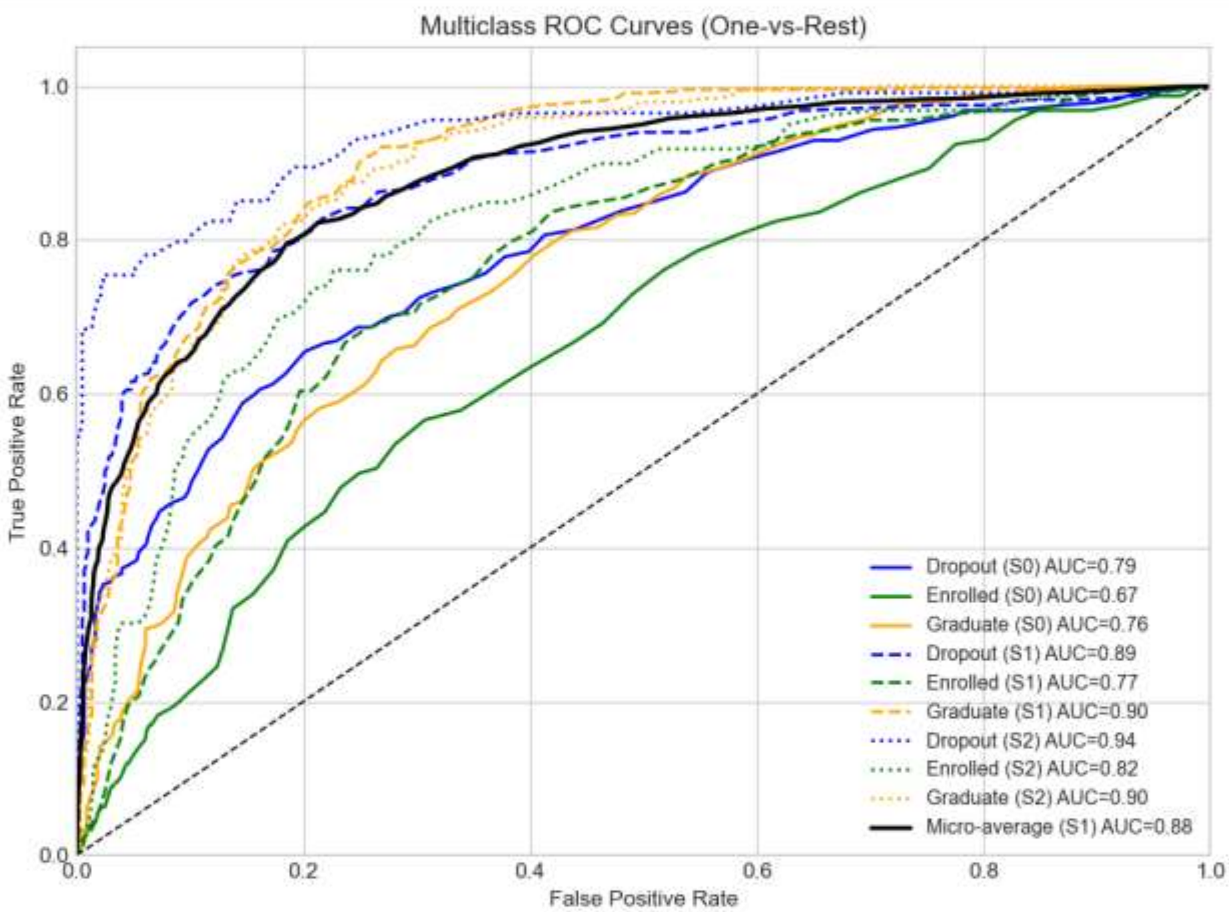

Figure.5 A multiclass ROC analysis using a one-vs-rest approach to evaluate the predictive performance

This figure 5 above presents a multiclass ROC analysis using a one-vs-rest approach to evaluate the predictive performance of models across three distinct phases (S0, S1, S2) for the classes "Dropout," "Enrolled," and "Graduate." The results demonstrate that model performance, as measured by AUC, generally improves from S0 to S1, with the highest AUC scores observed at the end of the first semester (S1), particularly for the "Graduate" class (AUC=0.90). Performance at S2 remains strong but shows slight degradation compared to S1, which is attributed to data attrition reducing the representation of the "Dropout" class. The micro-average AUC for S1 is 0.88, indicating robust overall classification capability when early academic data is available. This underscores the critical value of the first-semester data for accurate, early risk assessment in higher education.

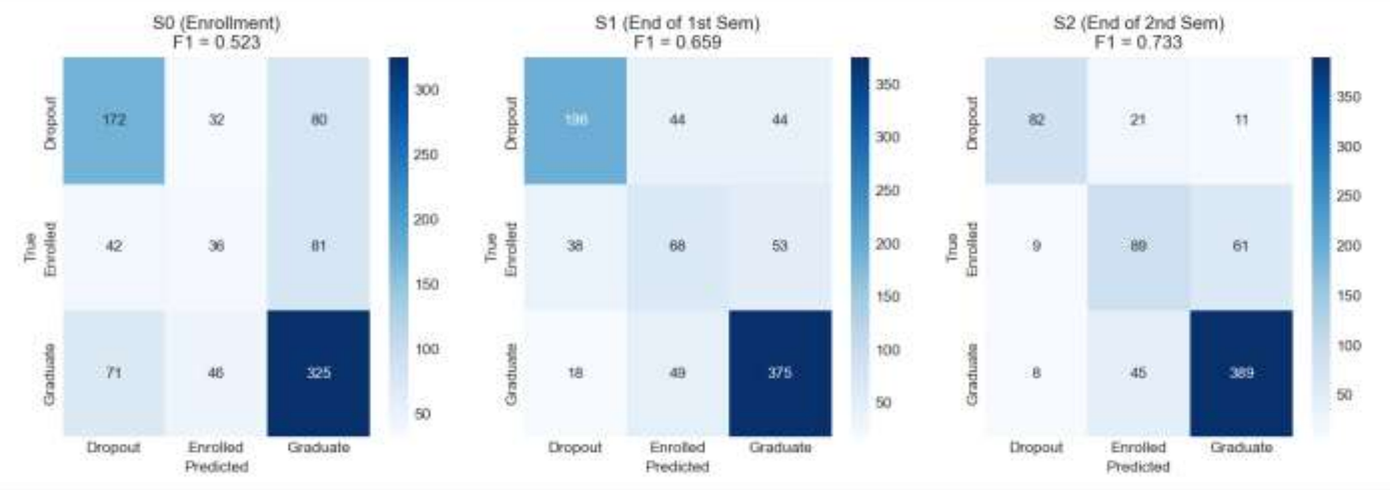

Figure 6 the confusion matrices for a multi-class student outcome prediction model across three temporal phases: enrollment (S0), end of the first semester (S1), and end of the second semester (S2).

The Figure 6 above results demonstrate a clear improvement in predictive performance over time, with the F1-score increasing from 0.523 at S0 to 0.659 at S1 and 0.733 at S2, indicating that early academic performance data is highly informative. The matrices reveal that while "Graduate" is consistently predicted with high accuracy, the minority classes "Dropout" and "Enrolled" are more challenging, particularly at S0. Notably, the number of actual "Dropout" cases decreases significantly by S2, likely due to student attrition, which impacts the model's ability to learn patterns for this class. This phased analysis underscores the optimal timing for intervention, suggesting that predictions made at the end of the first semester offer the best balance of accuracy and actionable insight for supporting at-risk students.

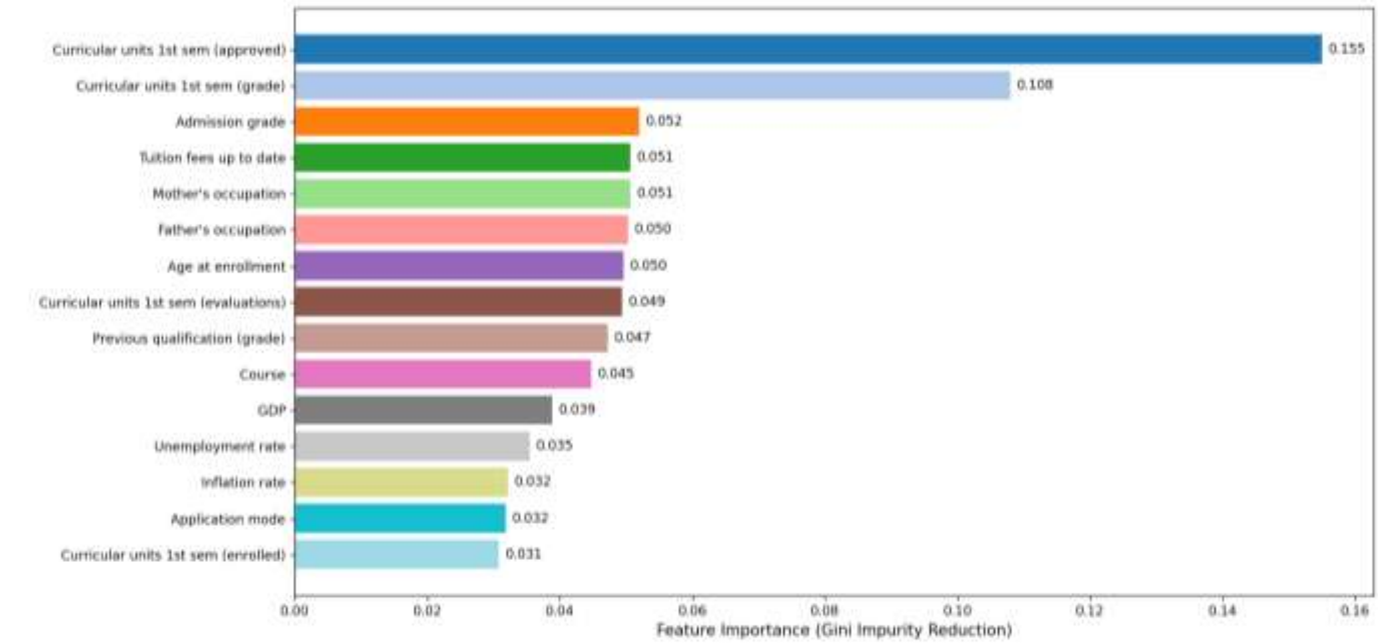

Figure. 7 the fifteen feature importance for predicting student outcomes

This bar chart displays the feature importance for predicting student outcomes, ranked by Gini Impurity Reduction. The most critical predictor is "Curricular units 1st sem (approved)," followed by "Curricular units 1st sem (grade)" and "Admission grade," indicating early academic performance is paramount. Socioeconomic factors like "Tuition fees up to date" and parental occupations also hold significant weight, while macroeconomic indicators like GDP and inflation are less influential. The results underscore that institutional and individual academic metrics are more predictive than broader economic conditions. This highlights the value of first-semester data for early intervention strategies in higher education.

4. Discussion

The reviewed studies collectively demonstrate that early prediction is feasible and valuable, with first-semester academic data being particularly informative. Ensemble methods, especially Random Forest variants, consistently outperform other classifiers in handling the dataset's complexity and imbalance. However, accuracy alone is insufficient. As Devic et al. (2024) argue, predictions used for high-stakes decisions (e.g., academic counseling, scholarship allocation) must also satisfy fairness and stability criteria. A model that performs well overall but discriminates against certain demographic or socioeconomic groups may exacerbate inequities. Moreover, Dyulicheva (2024) reminds us that predictive models are only one component of a larger learning analytics ecosystem. To truly improve academic qualification skills, institutions must:

The body of research surveyed in this review converges on a compelling insight: student academic trajectories in higher education are not predetermined but are highly responsive to timely, data-informed support (Dalla and Ahmad, 2023). As demonstrated by Realinho et al. (2022) and further refined by Martins et al. (2023), predictive models trained on multidimensional student data spanning demographic background, socioeconomic status, enrollment characteristics, and crucially, early academic performance can reliably flag those at risk of non-completion well before dropout becomes inevitable. Notably, ensemble methods such as Random Forest and its balanced variants consistently outperform other classifiers, particularly when explicitly designed or adapted to handle the pronounced class imbalance inherent in real-world educational datasets (Graduate: 50%, Dropout: 32%, Enrolled: 18%) (Lux et al., 2024). This robustness underscores the suitability of tree-based models for capturing the nonlinear, interactive nature of factors influencing student persistence.

However, the value of these models extends beyond raw predictive power. The phased prediction framework introduced by Martins et al. (2023) offers a critical operational insight: the optimal window for intervention is not at the point of enrollment when only static background variables are available but at the close of the first semester, once behavioral signals of academic engagement emerge. At this juncture, features such as the number of curricular units approved, semester grade averages, and attendance-related metrics become dominant predictors, reflecting a student's actual adaptation to the rigors of higher education (Jain et al., 2025). This finding reframes early warning systems not as static risk assessments but as dynamic, temporally sensitive instruments that align with pedagogical decision points. The slight performance decline observed in the second-semester phase (S2) further reveals a pragmatic reality: as students drop out, their records are often removed from institutional databases, thereby eroding the representation of the very population models aim to identify. This attrition effect highlights a key limitation of retrospective modeling and emphasizes the necessity of capturing data before disengagement becomes irreversible.

Beyond accuracy and timing, the ethical dimensions of predictive analytics in education cannot be overlooked. As Devic et al. (2024) compellingly argue, models deployed in high-stakes contexts such as academic counseling, scholarship allocation, or mentoring referrals must satisfy not only performance benchmarks but also fairness and stability criteria (Choi et al., 2025). Their Uncertainty-Aware (UA) ranking framework demonstrates that deterministic ranking systems, while intuitively appealing, are inherently fragile: minor perturbations in input predictions can yield drastically different outcomes, potentially amplifying systemic biases. In contrast, randomized, uncertainty-sensitive approaches preserve multigroup fairness when coupled with multiaccurate or multicalibrated predictors (Dalla, 2020). Although not yet applied directly to the Portuguese dropout dataset, this line of research provides a crucial normative foundation for future implementations especially in diverse student populations where algorithmic fairness is not a technical add-on but a pedagogical imperative (Dalla, 2020); (Ben Dalla et al., 2024). Moreover, Dyulicheva (2024) cautions against an overemphasis on student-level prediction at the expense of broader institutional analytics. Her critique that most educational datasets focus narrowly on learner behavior while neglecting instructor needs, curriculum design, and strategic policy resonates deeply with the limitations of even the most sophisticated student-risk models (Karczmarek et al., 2024). Predictive tools, no matter how accurate, remain siloed interventions unless embedded within a holistic learning analytics ecosystem (Achterberg et al., 2025). True enhancement of academic qualification skills requires more than identifying at-risk students; it demands actionable feedback loops that inform teaching practices (e.g., through instructor dashboards like CADA or SNAPP), adaptive curriculum adjustments, and resource allocation aligned with cohort-level trends. Taken together, these insights point toward an integrated model of support: one that leverages phased, interpretable, and fairness-aware prediction to trigger timely interventions personalized tutoring, financial aid, mental health resources while simultaneously feeding aggregated insights upward to inform systemic improvements (Ridwan et al., 2024); (Karczmarek et al., 2023); (Tito et al., 2023); (Tang et al., 2024); (Wang et al., 2023); (Santos et al., 2024). Achieving this vision, however, requires more than algorithmic refinement. It necessitates institutional commitment to data interoperability, staff training in data literacy, transparent communication with students about how their data is used, and continuous evaluation of real-world impact on retention and graduation rates not just model metrics. The ultimate measure of success for predictive modeling in higher education is not F1-score or balanced accuracy alone, but whether these tools genuinely empower students to develop the self-regulation, resilience, and metacognitive strategies that constitute robust academic qualification skills. The datasets and methods reviewed here provide a strong foundation but their transformative potential will only be realized when technical sophistication is coupled with pedagogical intentionality and ethical vigilance.

5. Conclusion

The landscape of predictive modeling in higher education has evolved substantially in recent years, driven by the convergence of rich, multidimensional student data and sophisticated machine learning methodologies. The Predict Students' Dropout and Academic Success dataset—curated from the Polytechnic Institute of Portalegre and encompassing over a decade of academic, demographic, and socioeconomic records has emerged as a cornerstone for empirical research in this domain. As demonstrated across multiple studies, this resource enables not only accurate forecasting of student trajectories but also actionable insights for timely pedagogical intervention. A central insight from this body of work is that effective prediction is not merely a function of algorithmic sophistication but of strategic data deployment. The end of the first semester constitutes an optimal window for risk assessment. At this juncture, early academic indicators such as approved curricular units and semester grades provide discriminative signals that significantly outperform models trained on enrollment-only data (S0) and avoid the attrition bias inherent in post-semester-two analyses (S2). This finding offers practical guidance for institutions aiming to balance predictive precision with real-time responsiveness in student support systems. Ensemble learning techniques particularly Random Forest and its gradient-boosted counterparts deliver robust baseline performance even in the face of pronounced class imbalance (Graduate: 50%, Dropout: 32%, Enrolled: 18%). Crucially, when paired with resampling or algorithmic rebalancing strategies (e.g., SMOTE, Balanced Random Forest), these models not only achieve high global F1-scores (up to 0.745) but also prioritize the accurate identification of at-risk students a group whose academic needs are most urgent yet statistically underrepresented.

However, predictive accuracy alone is insufficient for responsible deployment in educational contexts. Uncertainty-Aware (UA) ranking framework demonstrates that probabilistic predictions can be translated into equitable interventions that remain resilient to minor fluctuations in input data a vital safeguard against arbitrary or biased decision-making. This ethical dimension must be embedded into model design from the outset, particularly in diverse student populations where algorithmic fairness is not an optional enhancement but a pedagogical imperative. The most accurate and fair predictive system operates in isolation unless integrated into a broader, multi-stakeholder learning analytics ecosystem. Student-level predictions must be coupled with instructor-facing dashboards, curriculum-level trend analyses, and strategic institutional planning to holistically enhance academic qualification skills. True educational impact arises not from isolated risk scores but from coordinated support structures that empower students, inform teaching practice, and align with institutional missions. Looking ahead, the field must prioritize four key directions: (1) integrating behavioral data from learning management systems (e.g., Moodle, Canvas) to capture engagement dynamics beyond grades; (2) developing intervention recommender systems that translate risk predictions into tailored academic or financial support; (3) creating multilingual, culturally contextualized datasets to ensure global applicability; and (4) rigorously evaluating the real-world impact of predictive tools on retention and graduation outcomes through quasi-experimental designs. Harmonizing predictive power with pedagogical purpose, technical rigor with ethical responsibility, and individualized insight with systemic action. Only then can learning analytics fulfill its promise: transforming data not into deterministic labels, but into opportunities for student growth, resilience, and academic qualification.

Future Directions

While current research has established robust foundations for the early prediction of student dropout and academic performance, several compelling avenues remain for advancing both the technical and pedagogical dimensions of learning analytics. The ultimate goal is not merely to forecast risk but to catalyze actionable, equitable, and context-sensitive support that fosters students' academic qualification skills defined as the capacity for self-regulated learning, critical inquiry, and adaptive problem-solving. First, the integration of behavioral data from Learning Management Systems (LMS) such as Moodle, Canvas, or proprietary institutional platforms holds significant untapped potential. Current models, including those developed by Realinho et al. (2022) and Martins et al. (2023), rely predominantly on administrative and demographic variables, supplemented by semester-end academic outcomes. However, LMS logs capture rich, fine-grained indicators of student engagement: login frequency, video-watching behavior, discussion forum participation, assignment submission timing, and resource access patterns. Incorporating these behavioral signals could enable earlier and more nuanced risk detection potentially identifying disengagement before it manifests in low grades or missed evaluations. As illustrated in studies on online and blended learning (Qiu et al., 2022; Zhao et al., 2021), such data substantially enhances model sensitivity, especially during the critical first weeks of a course. Second, predictive systems must evolve from risk-scoring tools into intervention recommender frameworks. Knowing that a student is “at risk” provides limited utility without guidance on how to respond. Future work should focus on mapping specific risk profiles e.g., financial stress combined with low first-semester approvals to tailored support pathways: academic tutoring, mental health counseling, financial aid referrals, or time-management workshops. Such systems would function less as diagnostic labels and more as dynamic decision aids for advisors, aligning with Dyulicheva’s (2024) vision of multi-stakeholder learning analytics that serve both learners and educators.

Third, the field must move toward culturally and linguistically diverse datasets that reflect the global landscape of higher education. The benchmark dataset from the Polytechnic Institute of Portalegre, while methodologically rigorous, represents a specific national and institutional context (Portugal). To ensure broad applicability and avoid ethnocentric bias, researchers should curate and share datasets from varied educational systems accounting for differences in grading norms, academic calendars, socio-economic structures, and cultural attitudes toward help-seeking. Multilingual data repositories, particularly those incorporating non-Western contexts, will be essential for developing models that are both technically robust and socially inclusive. Finally, and perhaps most critically, the real-world impact of predictive analytics on student success must be empirically validated through quasi-experimental or longitudinal studies. Current evaluations rely heavily on classification metrics (e.g., F1-score, balanced accuracy), which, while useful for model comparison, do not measure whether interventions actually improve retention or graduation rates. Future research should partner with institutions to implement controlled trials comparing cohorts with and without LA-informed support to assess causal effects on academic trajectories. This shift from algorithmic performance to educational outcome is essential for establishing predictive analytics as a legitimate component of evidence-based pedagogical practice. By harmonizing advances in data richness, intervention intelligence, cultural adaptability, and impact evaluation, the next generation of learning analytics can transcend prediction to become a catalyst for student agency and academic flourishing. When fairness, interpretability, and pedagogical relevance are placed at the core of model design, institutions can transform data not into deterministic forecasts, but into empowering opportunities for growth, resilience, and lifelong qualification.

The Author's Acknowledgement

The authors gratefully acknowledge the use of artificial intelligence (AI) tools for proofreading and paraphrasing assistance during the preparation of this manuscript. As English is not the authors’ native language, AI-based language models were employed to improve grammatical accuracy, enhance clarity, and ensure adherence to academic writing conventions. However, all technical content, methodology, analysis, and conclusions presented in this work are the sole responsibility of the authors and reflect their original research and intellectual contribution.

References

Achterberg, J., Haas, M., van Dijk, B., & Spruit, M. (2025). Fidelity-agnostic synthetic data generation improves utility while retaining privacy. Patterns.

Adu-Twum, H. T., Sarfo, E. A., Nartey, E., Adesola Adetunji, A., Ayannusi, A. O., & Walugembe, T. A. (2024). Role of advanced data analytics in higher education: Using machine learning models to predict student success. Journal of Data Science and Artificial Intelligence, 3(1), 8.

Agila, A. A. A. (2024). Diabetes Prediction Using a Support Vector Machine (SVM) and visualize the results by using the K-means algorithm* Corresponding author:* Llahm Omar Faraj Ben Dalla. *Tarik Milod Alarbi Ahmad*, *2*.

Andreatos, A., & Leros, A. (2023). Can Oral Grades Predict Final Examination Scores? Case Study in a Higher Education Military Academy. Analytics, 2(4), 836-852.

Arora, D. (2023). Predicting Students Academic Success and Dropout Using Supervised Machine Learning. International Journal of Scientific Study, 11(6), 72-78.

Arwade, G., & Olafsson, S. (2025). Learning ensembles of interpretable simple structure. Annals of Operations Research, 1-29.

Beseiso, M. (2025). Enhancing student success prediction: A comparative analysis of machine learning technique. TechTrends, 69(2), 372-384.

Ben Dalla L., Medeni T.D., Medeni I.T., Ulubay M. (2024). Enhancing Healthcare Efficiency at Almasara Hospital: Distributed Data Analysis and Patient Risk Management. Economy: strategy and practice. 2024;19(4):54-72. https://doi.org/10.51176/1997-9967-2024-4-54-72

Choi, J., Karumbaiah, S., & Matayoshi, J. (2025). Bias or Insufficient Sample Size? Improving Reliable Estimation of Algorithmic Bias for Minority Groups. In Proceedings of the 15th International Learning Analytics and Knowledge Conference (pp. 547-557).

Dalla, L. O. B., Karal, Ö., & Degirmenciyi, A. (2025). Leveraging LSTM for Adaptive Intrusion Detection in IoT Networks: A Case Study on the RT-IoT2022 Dataset implemented On CPU Computer Device Machine.

Dalla, L. O. F. B. (2020). E-mail: mohmdaesed@ gmail. com E-mail: selflanser@ gmail. com Phone:+ 218945780716.

Dalla, L. O. B., Karal, Ö., Degirmenci, A., & Ali, M. يحصلا فرصلا هايم ةجلاعم تايلمع نيسحت تايمزراوخ لوح ةلماش تايبدلأل(2025). ةعجارم .

Dalla, L. O. F. B., & Ahmad, T. M. A. (2023). Heart Disease Prediction Via Using Machine Learning Techniques with Distributed System and Weka Visualization. *Journal of Southwest Jiaotong University*, *58*(4), 322-333.

Devic, S., Korolova, A., Kempe, D., & Sharan, V. (2024). Stability and Multigroup Fairness in Ranking with Uncertain Predictions. arXiv preprint arXiv:2402.09326.

Dalla, L. O. F. B. (2020). IT security Cloud Computing.

Devic, S., Korolova, A., Kempe, D., & Sharan, V. (2024). Stability and multigroup fairness in ranking with uncertain predictions. arXiv preprint arXiv:2402.09326.

Donnelly, J., Katta, S., Borgonovo, E., & Rudin, C. (2025). Doctor Rashomon and the UNIVERSE of Madness: Variable Importance with Unobserved Confounding and the Rashomon Effect. arXiv preprint arXiv:2510.12734.

Dyulicheva, Y. Y. (2024). Применение учебной аналитики в высшем образовании: датасеты, методы и инструменты. Высшее образование в России, 33(5), 86–111.

Erickson, N., Purucker, L., Tschalzev, A., Holzmüller, D., Desai, P. M., Salinas, D., & Hutter, F. (2025). Tabarena: A living benchmark for machine learning on tabular data. arXiv preprint arXiv:2506.16791.

https://archive.ics.uci.edu/dataset/697/predict+students+dropout+and+academic+success

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.

Jain, A., Dubey, A. K., Khan, S., Panwar, A., Alkhatib, M., & Alshahrani, A. M. (2025). A PSO weighted ensemble framework with SMOTE balancing for student dropout prediction in smart education systems. Scientific Reports, 15(1), 17463.

Karczmarek, P., Dolecki, M., Powroznik, P., Łagodowski, Z. A., Gregosiewicz, A., Gałka, Ł., ... & Jonak, K. (2023). Quadrature-inspired generalized Choquet integral in an application to classification problems. IEEE Access, 11, 124676-124689.

Karal, Ö., & Dalla, L. O. F. B. (2025). Lung Nodule Characterization in CT Scans Using Hybrid 3D Attention U-Net Segmentation and Transfer Learning-Based Classification Approach.

Karczmarek, P., Gregosiewicz, A., Łagodowski, Z. A., Dolecki, M., Gałka, Ł., Powroźnik, P., ... & Jonak, K. (2024). Analysis of smooth and enhanced smooth quadrature-inspired generalized Choquet integral. Fuzzy Sets and Systems, 483, 108926.

Kouloumpris, E., & Vlahavas, I. (2025). Markowitz random forest: Weighting classification and regression trees with modern portfolio theory. Neurocomputing, 620, 129191.

Li, W., Zhang, M., Zhang, L., & Liu, J. (2023). Integrated Private Data Trading Systems for Data Marketplaces. arXiv preprint arXiv:2307.16317.

Li, Y. Z., Ng, K. H., Khor, K. C., & Lim, Y. H. (2024). Predicting Undergraduate Academic Success with Machine Learning Approaches. In International Conference on Soft Computing and Data Mining (pp. 144-153). Cham: Springer Nature Switzerland.

Lux, N., Mandal, C., Decker, J., Luboeinski, J., Drewljau, J., Meisel, T., ... & Kunkel, J. (2024). HPC-AI benchmarks-a comparative overview of high-performance computing hardware and AI benchmarks across domains. J. Artif. Intell. Robot., 1, 2.

Martins, M. V., Baptista, L., Machado, J., & Realinho, V. (2023). Multi-Class Phased Prediction of Academic Performance and Dropout in Higher Education. Applied Sciences, 13(8), 4702.

Martins, M. V., Baptista, L., Machado, J., & Realinho, V. (2023). Multi-class phased prediction of academic performance and dropout in higher education. Applied Sciences, 13(8), 4702.

Martins, M. V., Tolledo, D., Machado, J., Baptista, L. M., & Realinho, V. (2021). Early prediction of student's performance in higher education: A case study. In World Conference on Information Systems and Technologies (pp. 166–175). Springer.

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.

Realinho, V., Machado, J., Baptista, L., & Martins, M. V. (2022). Predicting Student Dropout and Academic Success. Data, 7(11), 146.

Realinho, V., Machado, J., Baptista, L., & Martins, M. V. (2022). Predicting student dropout and academic success. Data, 7(11), 146.

Ridwan, A., Priyatno, A. M., & Ningsih, L. (2024). Predict Students' Dropout and Academic Success with XGBoost. Journal of Education and Computer Applications, 1(2), 1-8.

Romero, S., & Liao, X. (2025). Statistical and machine learning models for predicting university dropout and scholarship impact. PloS one, 20(6), e0325047.

Santos, S. O., Skiarski, A., García-Núñez, D., Lazzarini, V., Moral, R. D. A., Galvan, E., ... & Nepomuceno, E. (2024). Green machine learning: Analysing the energy efficiency of machine learning models. In 2024 35th Irish Signals and Systems Conference (ISSC) (pp. 1-6). IEEE.

Tang, Z., Jain, A., & Colina, F. E. (2024). A Comparative Study of Machine Learning Techniques for College Student Success Prediction. Journal of Higher Education Theory & Practice, 24(1).

Tang, Z., von Seekamm, K., Colina, F. E., & Chen, L. (2025). Enhancing Student Retention with Machine Learning: A Data-Driven Approach to Predicting College Student Persistence. Journal of College Student Retention: Research, Theory & Practice, 15210251251336372.

Tito, A. E. A., Condori, B. O. H., & Vera, Y. P. (2023). Análisis comparativo de Técnicas de Machine Learning para la predicción de casos de deserción universitaria. RISTI-Revista Ibérica de Sistemas e Tecnologias de Informação, (51), 84-98.

Villar, A., & de Andrade, C. R. V. (2024). Supervised machine learning algorithms for predicting student dropout and academic success: a comparative study. Discover Artificial Intelligence, 4(1), 2.

Wang, H., Li, J., & Zhu, G. (2023). A data feature extraction method based on the NOTEARS causal inference algorithm. Applied Sciences, 13(14), 8438.