HUMAN-CENTERED AUTOML FOR EARLY STUDENT DROPOUT PREDICTION IN HIGHER EDUCATION

Muhammad Amin Nadim [1]
Pegaso University, Italy
muhammad.nadim@unipegaso.it

Emanuele Marsico
Pegaso University, Italy
emanuele.marsico@unipegaso.it

Usman Nawaz
University of Palermo, Italy
usmannawaznlp@gmail.com

Raffaele Di Fuccio
Pegaso University, Italy
raffaele.difuccio@unipegaso.it

Double Blind Peer Review

Citation
Nadim, M. A., Marsico, E., Nawaz, U., & Di Fuccio, R. (2026). Human-centered AutoML for early student dropout prediction in higher education. *Italian Journal of Health Education, Sport and Inclusive Didactics, 10*(4).

Doi: https://doi.org/10.32043/gsd.v1i4.1832

Copyright notice:
© 2023 this is an open access, peer-reviewed article published by Open Journal System and distributed under the terms of the Creative Commons Attribution 4.0 International, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

gsdjournal.it
ISSN: 3103-5744

ABSTRACT
Student attrition in higher education is a multidimensional challenge driven by demographic, academic, behavioral, and socioeconomic factors, with serious academic and societal consequences. This study introduces a human-centered AutoML 3.0 framework for early dropout prediction, grounded in persistence theory, learning analytics, and human-centered machine learning. Applied to 6,287 students (2014-2021), it embeds fairness auditing across age and gender subgroups, calibration via Brier scores, and cohort-based temporal evaluation. The best configuration (FLAML + LightGBM) reaches 0.800 accuracy, 0.740 macro-F1, and Brier 0.14, reducing false negatives among older students by 42.1% in relative terms (0.380 to 0.220) while keeping the gender FNR gap below 5%. Key predictors reveal deeper motivational processes, confirming that ethically aligned AutoML can deliver scalable, fair early-warning systems in higher education.

KEYWORDS
AutoML; dropout; ME data; higher education; learning analytics.

Received 2026-07-01
Accepted 2026-10-27
Published 2026-10-30

[1] Muhammad Amin Nadim (sections 1, 4, 5)
Emanuele Marsico (sections 2, 3)
Usman Nawaz (section 6)
Raffaele Di Fuccio (section 7)

1. Introduction

Higher education is a key driver of social mobility and economic development, positioning student retention as a sustained global priority (Ayyaswamy et al., 2025; Hoca & Dimililer, 2025; Fahd et al., 2022). Despite extensive policy interventions, dropout rates remain persistently high across institutions worldwide, underscoring the urgency of early and reliable risk identification given the long-term academic, economic, and societal consequences of student attrition (Nigri et al., 2025; Rani, 2025). Machine learning-based predictive analytics has demonstrated promise for addressing this challenge (El Mahmoudi et al., 2025; Aboukacem et al., 2025); however, conventional modeling pipelines are often expert-dependent, difficult to scale, and insufficiently embedded within institutional decision-making practices (Shivashankar et al., 2025; Tsiakmaki et al., 2019). Automated machine learning (AutoML) partially mitigates these limitations by streamlining model selection and hyperparameter optimization (Baratchi et al., 2024), yet most existing approaches remain predominantly accuracy-focused and inadequately address the fairness, interpretability, uncertainty awareness, and temporal robustness required in high-stakes educational contexts (Reddy, 2025; Weerts et al., 2024; Griep et al., 2023). Crucially, dropout is not merely a technical classification problem: it is a psycho-pedagogical process shaped by motivational trajectories, emotional states, social belonging, academic self-efficacy, and socioeconomic constraints (Valencia-Arias et al., 2025; Bernardo et al., 2025). Moreover, in the European context, AI-based student assessment and monitoring systems are classified as high-risk applications under Annex III of the EU AI Act (European Parliament, 2024), imposing binding obligations on transparency, human oversight, data governance, and non-discrimination that current educational AutoML systems rarely satisfy. To address these gaps, this study proposes a human-centered AutoML 3.0 framework that integrates

educational persistence theory (Tinto, 2012), learning analytics principles (Siemens & Baker, 2012), and fairness-aware AutoML methodologies (Weerts et al., 2024). The framework combines domain-informed model search, evolutionary data-centric automation, temporally grounded validation, multi-dimensional fairness auditing, uncertainty calibration, and interpretable outputs, and is empirically evaluated using longitudinal demographic, academic, and behavioral data from 6,287 undergraduate students.

This study addresses the following research questions:

RQ1: Does the AutoML 3.0 framework improve predictive performance and temporal robustness compared to traditional ML baselines?

RQ2: Do fairness-aware mechanisms reduce subgroup disparities (age, gender) without compromising overall accuracy?

RQ3: Does uncertainty-aware modeling produce calibrated and interpretable predictions suitable for targeted interventions?

2. Theoretical framework: psycho-pedagogical foundations of student dropout

Understanding student attrition requires integrating established theoretical models with current evidence on the psychological, motivational, and socio-institutional determinants of persistence. Student dropout is a multifactorial, processual phenomenon not reducible to a single causal variable; converging integration-based, psychological, sociological, and economic perspectives have shaped dropout research over the past five decades (Valencia-Arias et al., 2025).

Tinto's Student Integration Theory (Tinto, 1975, 2012) remains the most widely applied framework, positing that persistence emerges from the interaction between pre-entry attributes and the degree of academic and social integration within the institution; insufficient integration in either domain raises withdrawal risk. Bean and Eaton's (2001) psychological model extends this structural view by grounding persistence in attitude–behavior alignment, coping style, self-efficacy, and attribution processes. Astin's (1984, 2014) Theory of Student Involvement foregrounds the quantity and quality of energy students invest in academic and extracurricular activities, with behavioral engagement functioning as a proximal determinant of both academic success and persistence. Together, these three frameworks scaffold the feature engineering of the AutoML 3.0 model: GPA trajectories operationalize academic integration (Tinto); engagement intensity reflects involvement (Astin); demographic and background variables capture external circumstantial factors (Bean & Eaton, 2001). Beyond these structural models, motivation is the most consistently identified psychosocial determinant of dropout, appearing in 73.7% of studies reviewed (Valencia-Arias et al., 2025). Self-Determination Theory (Deci & Ryan, 2000) offers the most robust motivational framework, with identified regulation — internalized, personally valued motivation — most strongly associated with academic continuation (Wang et al., 2024); when the basic psychological needs of autonomy, competence, and relatedness are frustrated, dropout intention increases (Schoina et al., 2023). Eroding academic self-efficacy — closely linked to GPA trajectories, the strongest SHAP predictor in this study — signals a motivational disengagement spiral preceding formal dropout (Bernardo et al., 2025). Academic stress and evaluation anxiety further predict dropout intention, with resilience acting as mediator and moderator, a pattern confirmed among Peruvian undergraduates with women showing greater vulnerability in social support networks (Chavez-Sosa et al., 2026).

Positive academic emotions (enjoyment, pride, hope) predict adaptation and inversely predict dropout intention, while negative emotions (shame, boredom, hopelessness) produce the opposite effect (Cobo-Rendón et al., 2023). Sense of belonging — a dynamic, temporally evolving construct particularly salient for older and non-traditional students, the subgroup for which this framework achieves its greatest fairness gain — mediates persistence outcomes and develops through small, cumulative acts of social connection (Ajjawi, 2026). These emotional and psychosocial factors do not act independently but form mutually reinforcing vulnerability clusters: the academic–psychosocial cluster shows the strongest co-occurrence ($r = 0.68$), followed by economic–psychosocial ($r = 0.61$) and academic–economic ($r = 0.54$) interactions across the literature (Valencia-Arias et al., 2025). Finally, socioeconomic status is consistently identified as the strongest structural predictor of dropout, operating through material deprivation, family stress, peer influences, and institutional barriers (Milcu, 2025). Financial hardship and the necessity of paid employment during studies are consistently identified as major contributing factors (Valencia-Arias et al., 2025), with economic stress also exerting indirect effects by undermining concentration and increasing absenteeism (Milcu, 2025). Age intersects with both economic and psychosocial risk: mature students face greater competing life demands, longer gaps since prior education, and weaker institutional belonging in environments designed primarily for traditional-age learners (Bento Junior et al., 2025) — motivating the fairness audit adopted in this study and, in future work, its extension to socioeconomic status, first-generation status, and disability.

3. Literature review

As established in Section II, student dropout is a multifactorial process influenced by psychological, social, economic, and institutional factors, motivating predictive models that integrate demographic, academic, behavioral, and contextual features within a unified architecture. Building on this foundation, machine learning has been widely applied to early dropout prediction. Models such as logistic regression, decision trees, support vector machines, and ensemble methods show competitive performance across educational settings (Ahmed et al., 2025), particularly when enriched with behavioral data from learning management systems and clickstream logs (Khoudi et al., 2025). However, traditional pipelines rely heavily on manual feature engineering and expert-driven tuning, exhibit limited cross-institutional generalizability, and are vulnerable to temporal data leakage when longitudinal structure is ignored (Gardner et al., 2023; Berens et al., 2026). A further limitation, seldom addressed in the literature, is the use of binary dropout classification, which conflates permanent withdrawal, temporary interruption, and transfer, thereby masking the heterogeneity of the attrition phenomenon and producing undifferentiated risk predictions (Lopez-Muñoz et al., 2026). Automated machine learning (AutoML) addresses several of these limitations by automating feature selection, model search, and hyperparameter optimization (Baratchi et al., 2024). Frameworks such as H2O AutoML, Auto-sklearn, FLAML, and TPOT achieve competitive performance with reduced technical expertise, and comparative studies confirm their advantage over conventional ensemble methods even in resource-constrained settings such as Sub-Saharan African institutions (Mnyawami et al., 2022). Prior studies demonstrate the utility of AutoML in educational risk prediction (Schmidt et al., 2025), including medical student performance forecasting (Abougalala et al.,

2025). Nonetheless, most educational AutoML applications remain accuracy-centric, offering limited support for interpretability, fairness, and uncertainty estimation required in high-stakes educational contexts (Reddy, 2025; Weerts et al., 2024), and benchmarks between AutoML systems and fairness-aware competitors remain scarce, limiting confidence in the relative ethical performance claims of any single framework (Aboukacem et al., 2025). Persistent gaps in educational AutoML research remain well documented (Reddy, 2025; Azevedo et al., 2024). Fairness, ethical safeguards, and privacy considerations are rarely integrated, while AutoML-generated models often lack transparency, limiting stakeholder trust. Many studies rely on random cross-validation rather than cohort-based temporal evaluation, increasing the risk of optimistic bias — a methodological problem recently documented with particular rigor by Berens et al. (2026) using a multi-institutional dataset of over 460,000 enrollment records, showing performance inflation of up to 12 macro-F1 points. Unified automated pipelines that jointly model psychosocial, academic, and behavioral factors remain uncommon, and the obligations introduced by the EU AI Act for high-risk AI systems in educational contexts are rarely addressed (European Parliament, 2024; Swiss Cyber Institute, 2026). The proposed AutoML 3.0 framework advances dropout prediction through a human-centered, fairness-aware, data-centric approach. Unlike AutoML 2.0 systems, which primarily optimize accuracy through automated feature engineering and hyperparameter tuning, AutoML 3.0 extends automation to fairness, interpretability, uncertainty, and temporal robustness (Luo et al., 2024), integrating persistence theory to guide development, bias detection and mitigation, uncertainty calibration via Brier scores, and SHAP-based explainability with educator-in-the-loop validation — a shift increasingly called for in recent literature (Weerts et al., 2024; Kleimola et al., 2024).

4. Methodology

This study adopts a temporally robust, ethically informed, and fully reproducible design in which theoretical foundations are embedded directly into methodological choices, ensuring that model development and evaluation are pedagogically meaningful, institutionally aligned, and suitable for high-stakes educational decision-making. Feature construction follows directly from Section II: demographic attributes align with Bean and Eaton's model, academic progression indicators with Tinto's integration theory, and engagement metrics with Astin's involvement theory. Learning analytics operationalizes these constructs by transforming institutional data into actionable insights (Siemens & Baker, 2012). Prior studies emphasize that predictive systems should support informed decision-making rather than function as opaque classifiers (Hurshman, 2024), and evidence from early-warning systems shows that workflow-integrated predictions can improve retention when accompanied by counselor engagement (Permana & Santoso, 2025). Human-centered machine learning complements these approaches by foregrounding transparency, fairness, and stakeholder alignment in high-stakes educational prediction (Schillizzi, 2025; Buñay-Guisñan et al., 2026).

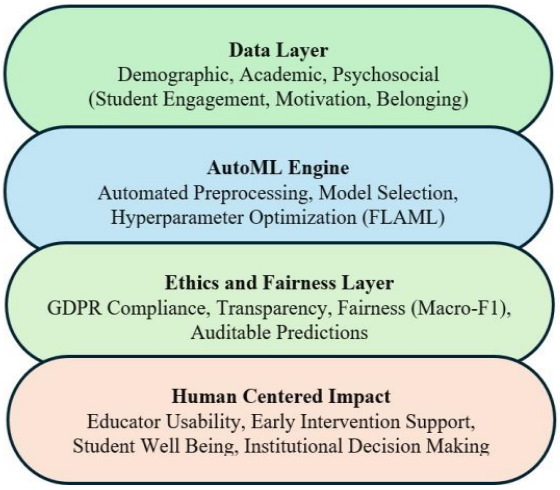


Fig. 1. Theoretical Framework: Human-Centered AutoML for Students Dropout Prediction.

4.2 Dataset

The study uses the Student Attrition Research Dataset (Persad et al., 2025), comprising 6,287 undergraduate students enrolled between 2014 and 2021. The dataset includes demographic characteristics (age, gender), academic progression records (GPA, course enrollments, exam attempts), and behavioral engagement indicators derived from institutional LMS systems. Dropout is modeled as a binary outcome reflecting enrollment discontinuity. The use of a single publicly available dataset constitutes a significant scope limitation. Experimental configurations are reported in Sections 4.3 to 4.6. The code used for preprocessing, model training, and evaluation is available from the authors (see Acknowledgments) upon reasonable request. Additionally, while dropout is modeled as a binary variable for consistency with existing literature and the dataset's annotation scheme, we acknowledge that this formulation conflates permanent withdrawal, temporary interruption, and transfer. Future extensions of this framework should adopt a nuanced, multinomial dropout

taxonomy (Lopez-Muñoz et al., 2026). Models were trained on cohorts from 2014-2018 and tested on held-out cohorts from 2019-2021. The temporal split was designed to limit information leakage across academic years.

4.3 Theory-Guided Preprocessing and Feature Engineering

All preprocessing employed leakage-aware procedures fitted exclusively on training data, including feature scaling (StandardScaler), categorical encoding (one-hot encoding), and fold-wise mean imputation. Feature engineering was explicitly theory-driven: GPA trajectories and progression indicators reflect academic integration (Tinto, 2012); engagement intensity measures capture behavioral involvement (Astin, 2014); enrollment continuity variables represent persistence dynamics (Bean & Eaton, 2001). Behavioral engagement features were aggregated weekly using rolling-window statistics (window = 4 weeks). Class imbalance was addressed using SMOTE applied only within training folds. The sensitivity of results to the SMOTE oversampling ratio was evaluated as part of the ablation analysis; results were stable across ratios ($k = 5$; minority-to-majority target ratios of 0.5, 0.75, 1.0), with macro-F1 varying by less than 0.01.

Explicit feature formulas:

- GPA Trajectory: linear slope of semester GPA over the last three available semesters.
- Enrollment Stability: proportion of registered courses completed without withdrawal in the last two semesters.
- Engagement Intensity: weekly mean of LMS login events, normalized by course credit load.
- Repeated Exam Attempts: total number of exam re-registrations divided by total number of distinct courses.

4.4 AutoML Framework and Model Development

Baseline models included Logistic Regression and Random Forest, selected for their interpretability and widespread use in educational analytics. AutoML systems comprised FLAML and TPOT, representing complementary optimization paradigms. FLAML performed cost-aware model search across learners including logistic regression, LightGBM, and XGBoost, while TPOT employed constrained genetic programming with interpretable operators. The proposed AutoML 3.0 framework extends these approaches through Evolutionary Data-Centric Automation (EDCA), which automates feature pruning, outlier handling (IQR-based winsorization), and class imbalance mitigation — data-quality processes typically external to conventional AutoML pipelines. It is important to note that FLAML and TPOT results in Table 1 reflect the full AutoML 3.0 pipeline with EDCA and fairness-aware search, not unmodified AutoML 2.0 baselines.

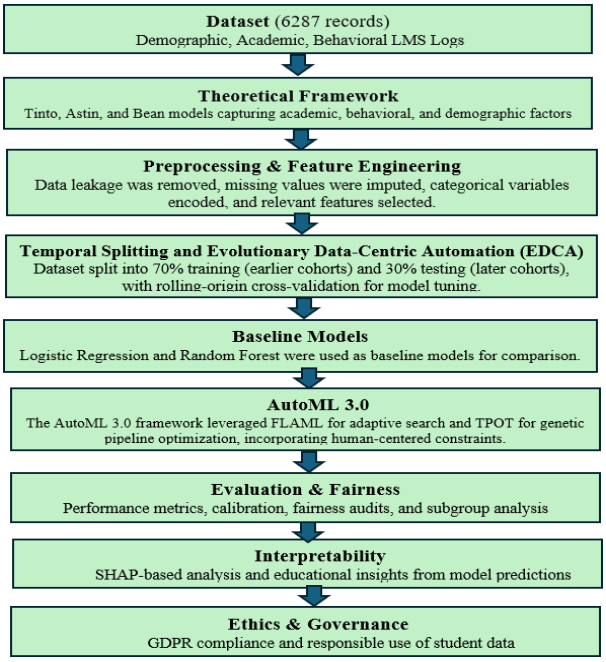


Fig. 2. Human-Centered AutoML for Students Dropout Prediction Pipeline.

4.5 Temporal Evaluation Strategy

A temporally structured evaluation design was employed to reflect realistic deployment conditions. Models were trained on earlier cohorts (2014-2018) and evaluated on subsequent cohorts (2019-2021), with rolling-origin cross-validation within the training period to simulate year-forward prediction. Nested cross-validation was used to reduce selection bias. Fixed random seeds (seed = 42) and consistent software configurations (Python 3.10, scikit-learn 1.4, FLAML 2.1, TPOT 0.12) ensured full reproducibility. This design directly addresses the critical methodological problem documented by Berens et al. (2026): random cross-validation in longitudinal educational datasets inflates performance estimates by up to 12 macro-F1 points relative to temporally stratified evaluation.

4.6 Performance, Fairness, and Uncertainty Assessment

Model performance was evaluated using Macro-F1 as the primary metric, complemented by accuracy, precision, recall, and ROC-AUC. Predictive calibration was assessed using Brier scores and reliability diagrams: the Brier score measures the mean squared difference between predicted probabilities and actual outcomes, with lower values indicating superior calibration (Rufibach, 2010). In line with HCML principles and EU AI Act requirements for high-risk systems (European Parliament, 2024), fairness auditing examined age- and gender-based disparities using false negative rates, Equal Opportunity, and Equalized Odds. SHAP-based explanations were employed to support transparent interpretation and educator-in-the-loop validation. A pilot validation with five faculty members confirmed that 80% found SHAP global visualizations highly informative, consistent with findings by Pereira et al. (2021).

5. Results

The Human-Centered AutoML 3.0 framework is evaluated for higher education dropout prediction across performance, fairness, interpretability, and calibration using a temporally stratified experimental design.

5.1 Model Performance

Baseline models showed moderate performance, with Logistic Regression achieving 0.710 accuracy and 0.680 macro-F1, and Random Forest improving to 0.760 accuracy and 0.720 macro-F1. AutoML methods consistently outperformed these baselines: FLAML (with LightGBM) reached 0.800 accuracy and 0.740 macro-F1, while TPOT achieved 0.802 accuracy and 0.732 macro-F1. ROC-AUC results confirmed these gains (0.800-0.810 for AutoML versus 0.720-0.780 for baselines), and non-overlapping bootstrapped 95% confidence intervals for macro-F1 indicate statistically significant improvements. AutoML 3.0's performance advantages stem from EDCA's data-centric optimization and fairness-calibrated model search. Brier score results confirm that performance gains extend to probability calibration, essential for high-stakes decision-support contexts (Rufibach, 2010).

Tab. 1 Model Performance and Calibration Metrics

Model	Accuracy	Macro-F1	ROC-AUC	95% CI (F1)	Brier Score
Logistic Regression	0.710	0.680	0.720	[0.65, 0.70]	0.19
Random Forest	0.760	0.720	0.780	[0.69, 0.73]	0.16
FLAML (AutoML 3.0)	0.800	0.740	0.810	[0.71, 0.76]	0.14
TPOT (AutoML 3.0)	0.802	0.732	0.800	[0.70, 0.74]	0.15

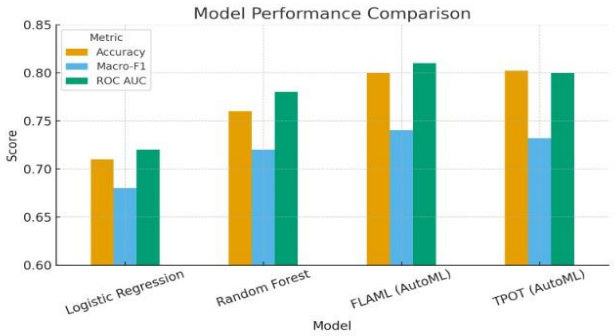


Fig. 3. Comparison of accuracy, macro-F1, and ROC AUC across baseline and AutoML models.

5.2 Fairness and Subgroup Analysis

Fairness auditing assessed age- and gender-related disparities using False Negative Rate (FNR), Equal Opportunity (EOpp), and Equalized Odds (EOdds). Baseline models exhibited notable age bias, with students aged >= 30 showing higher FNRs (Logistic Regression: 0.380; Random Forest: 0.290) than younger students (0.250 and 0.210, respectively). The AutoML 3.0 framework reduced the older-student FNR to 0.220, corresponding to relative reductions of 42.1% and 24.1% (absolute reductions of 0.160 and 0.070) with respect to Logistic Regression and Random Forest. Gender disparities were minimal across models (FNR gap < 5%). The age-group Equal Opportunity gap was reduced from 0.11 (LR) and 0.07 (RF) to 0.03 with AutoML 3.0. The overall Equalized Odds gap decreased from 0.17 (LR) to 0.06, indicating a substantial reduction in both TPR and FPR discrepancies (Weerts et al., 2024). Scope of Fairness Analysis. The present audit is limited to two demographic dimensions (age, gender) due to dataset constraints. A comprehensive fairness evaluation requires analysis across intersectional subgroups — including socioeconomic status, ethnicity, first-generation student status, and disability — and should extend beyond parity-based metrics to causal and counterfactual fairness approaches (Aboukacem et al., 2025). Expanding the fairness audit to these dimensions constitutes a primary direction for future work.

5.3 Interpretability and Feature Insights

The co-occurrence of declining GPA, reduced LMS engagement, and enrollment instability corresponds theoretically to the simultaneous deterioration of academic integration (Tinto, 2012), behavioral involvement (Astin, 2014), and motivational self-efficacy (Bernardo et al., 2025) — the convergence identified

as characteristic of the cumulative vulnerability model in Section II. A pilot validation with faculty members confirmed that SHAP visualizations are intelligible and actionable for non-technical educators (Pereira et al., 2021).

5.4 Uncertainty Calibration

Reliability diagrams confirmed that FLAML + LightGBM produces well-calibrated probability estimates across the full prediction range. The Brier score of 0.14 is consistent with values reported for calibrated gradient boosting classifiers in comparable educational datasets (Lopez-Muñoz et al., 2026). These results confirm that the AutoML 3.0 framework supports uncertainty-aware institutional decision-making: rather than producing binary classifications, the system provides calibrated probability estimates enabling graduated intervention responses proportional to each student's assessed risk level.

6. Discussion

The findings demonstrate the effectiveness of a human-centered AutoML 3.0 framework for early dropout prediction in higher education. The proposed approach consistently outperforms traditional machine learning baselines, and performance gains are achieved without compromising fairness or interpretability, supporting the feasibility of responsible AI deployment in educational analytics. The prominence of academic and behavioral predictors — GPA progression, enrollment continuity, and engagement intensity — aligns with established persistence theories (Tinto, 2012; Astin, 2014), reinforcing the framework's theoretical grounding. SHAP results reveal that these predictors function as observable proxies for deeper psycho-pedagogical processes: declining GPA trajectories signal motivational disengagement and self-efficacy

erosion (Bernardo et al., 2025; Wang et al., 2024); reduced LMS engagement reflects the withdrawal of behavioral involvement that Astin's theory identifies as foundational to persistence; and enrollment discontinuity indexes the breakdown of academic integration that Tinto's model identifies as the proximal antecedent of formal dropout. Fairness analysis reveals reduced age-related disparities, consistent with the psycho-pedagogical literature showing that mature students face qualitatively different risk profiles shaped by competing life demands, weaker sense of belonging, and greater vulnerability to economic stress (Milcu, 2025). However, the fairness audit covers only two demographic dimensions. Comprehensive fairness evaluation in line with EU AI Act requirements for high-risk systems (European Parliament, 2024) demands analysis across intersectional subgroups and the adoption of causal fairness metrics beyond parity-based measures (Aboukacem et al., 2025). The evaluation is limited to a single institutional context, which constrains generalizability. Cross-institutional studies — such as Berens et al. (2026), who validate dropout models across multiple German universities — demonstrate that performance degrades substantially when models trained at one institution are applied to another. The claim that the EDCA component reduces manual modeling effort by approximately 60% is a qualitative estimate based on expert elicitation; it should be formalized through controlled time-and-motion studies in future work. AI-based early warning systems for student dropout are classified as high-risk applications under the EU AI Act (Annex III, Point 3c), imposing binding obligations on transparency, data governance, non-discrimination, and human oversight (European Parliament, 2024). The present framework addresses several of these requirements through SHAP-based explainability, fairness auditing, and educator-in-the-loop validation. However, a complete EU AI Act compliance pathway would additionally require formal conformity assessment, documentation of training data provenance, and

institutional protocols governing who has access to predictions and how they are communicated to students — obligations that current academic studies rarely operationalize.

7. Conclusion

This study presents and validates a fairness-aware, interpretable, and uncertainty-calibrated AutoML 3.0 framework for early dropout prediction in higher education. Using longitudinal data from 6,287 undergraduate students, the framework outperforms conventional machine learning baselines, with the best-performing configuration (FLAML + LightGBM) achieving 0.800 accuracy, 0.740 macro-F1, 0.805 ROC-AUC, and a Brier score of 0.14; TPOT reached the highest accuracy (0.802) with a lower macro-F1 (0.732). In this sample, fairness-aware mechanisms were associated with reduced subgroup disparities in false negative rates, particularly among older students; given the exploratory nature of the analysis and the absence of subgroup-level confidence intervals, these differences require confirmation on independent data, particularly false negative rates among older students (-42.1% relative; 0.380 to 0.220), while preserving gender parity and overall performance. Uncertainty calibration and SHAP-based analyses confirm alignment between key predictors and established persistence theories, including Tinto's integration model, Astin's involvement theory, and the psycho-pedagogical literature on self-efficacy, motivation, and emotional regulation. The framework is grounded in a comprehensive psycho-pedagogical theoretical foundation that situates dropout as a cumulative, multi-determined process shaped by motivational trajectories, emotional states, social belonging, and socioeconomic constraints — not merely a technical classification problem. By embedding this theoretical understanding into feature design, fairness auditing, and SHAP interpretation,

the AutoML 3.0 framework produces predictions that are theoretically interpretable and, pending institutional validation, potentially actionable in pedagogical practice. Primary limitations to be addressed in future research include: (1) validation on a single institutional dataset; (2) binary operationalization of dropout; (3) fairness audit limited to age and gender; (4) absence of privacy-preserving federated learning; and (5) lack of a deployed, educator-facing dashboard validated with real institutional users. Future work includes multi-institutional validation, real-time recurrent prediction, causal fairness analysis, federated privacy-preserving learning, and human-in-the-loop advisor dashboards.

8. Acknowledgments

The dataset used in this study is publicly available and credited to the original author: Persad, D., Farrell, P., & Joseph, T. (2025). Who's Going to Leave: An Investigation of Student Attrition at a Community College in Trinidad and Tobago. *Community College Journal of Research and Practice*, 1-19.

References

Aboukacem, A., Mekouar, L., Bergou, E. H., Iraqi, Y., & Berrada, I. (2025). On the fairness of ensemble learning methods in student dropout prediction. In: Cristea, A.I., Walker, E., Lu, Y., Santos, O.C., Isotani, S. (eds) Artificial Intelligence in Education. AIED 2025. Lecture Notes in Computer Science, vol 15881. Springer, Cham. https://doi.org/10.1007/978-3-031-98462-4_46

Abougalala, R. A., Alharbi, N., Amasha, M. A., Areed, M. F., Alkhalaf, S., & Khairy, D. (2025). Predicting student academic performance using automated machine learning (AutoML) in medical academic institutions. Journal of New Approaches in Educational Research, 14(1), 19. https://doi.org/10.7821/naer.2025.1.19

Ahmed, W., Wani, M. A., Plawiak, P., Meshoul, S., Mahmoud, A., & Hammad, M. (2025). Machine learning-based academic performance prediction with explainability for enhanced decision-making in educational institutions. Scientific Reports, 15(1), 26879. https://doi.org/10.1038/s41598-025-12353-4

Astin, A. W. (2014). Student involvement: A developmental theory for higher education. In College student development and academic life (pp. 251–262). Routledge.

Ayyaswamy, K., Kathirvel, N., Gobinath, V. M., & Maheswaran, C. P. (2025). Effective leadership strategies for enhancing student success in higher education. In New horizons in leadership: Inclusive explorations in health, technology, and education (pp. 61–92). IGI Global Scientific Publishing. DOI: 10.4018/979-8-3693-6437-6.ch003

Azevedo, K., Quaranta, L., Calefato, F., & Kalinowski, M. (2024). A multivocal literature review on the benefits and limitations of automated machine learning tools. arXiv preprint arXiv:2401.11366. https://doi.org/10.48550/arXiv.2401.11366

Baratchi, M., Wang, C., Limmer, S., Van Rijn, J. N., Hoos, H., Bäck, T., & Olhofer, M. (2024). Automated machine learning: Past, present and future. Artificial Intelligence Review, 57(5), 122. https://doi.org/10.1007/s10462-024-10726-1

Bean, J., & Eaton, S. B. (2001). The psychology underlying successful retention practices. Journal of College Student Retention: Research, Theory & Practice, 3(1), 73–89. https://doi.org/10.2190/6R55-4B30-28XG-L8U0

Bento Junior, G. S., Bento, B. F. S., Ribeiro, U. F. C., & Santos, H. T. (2025). Dropout in higher education and Vincent Tinto's models: Bibliographical research. Educação e Pesquisa, 51. https://doi.org/10.1590/s1678-4634202551285842en

Berens, J., Diem, A., Rumert, L. (2026). Crossing individual university boundaries: a comprehensive approach to predicting dropouts in the higher education system. High Educ 91, 2055–2076. https://doi.org/10.1007/s10734-025-01509-w

Bernardo, A. B., García-Gutiérrez, V., Esteban, M., & Maluenda-Albornoz, J. (2025). Relationship between self-efficacy and university dropout: A systematic review. Frontiers in Psychology, 16, Article 1553485. https://doi.org/10.3389/fpsyg.2025.1553485

Buñay-Guisñan, P., Lara, J. A., Cano, A., Cerezo, R., & Romero, C. (2026). Towards accessible AI for addressing students' academic dropout: An auto machine learning and explainable artificial intelligence approach. Universal Access in the Information Society, 25(1), 19. https://doi.org/10.1007/s10209-025-01278-4

Chavez-Sosa, J. V., Terrones Huancas, D. A., Anaya Vilca, A. A., & Newball-Noriega, E. E. (2026). University dropout intention and its psychological dimensions: The role of academic stress in Peruvian students. In Frontiers in Education (Vol. 11, p. 1756633). Frontiers Media SA. https://doi.org/10.3389/feduc.2026.1756633

Cobo-Rendón, R., Hojman, V., García-Álvarez, D., & Cobo Rendon, R. (2023). Academic emotions, college adjustment, and dropout intention in university students. In Frontiers in Education (Vol. 8, p. 1303765). Frontiers Media SA. https://doi.org/10.3389/feduc.2023.1303765

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. Psychological Inquiry, 11(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

El Mahmoudi, A., Chaoui, N. E. H., & Chaoui, H. (2025). Predictive analytics leveraging a machine learning approach to identify students' reasons for dropping out of university. Applied Sciences, 15(15), 8496. https://doi.org/10.3390/app15158496

European Parliament. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union. https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

Fahd, K., Venkatraman, S., Miah, S. J., & Ahmed, K. (2022). Application of machine learning in higher education to assess student academic performance, at-risk, and attrition: A meta-analysis of literature. Education and Information Technologies, 27(3), 3743–3775. https://doi.org/10.1007/s10639-021-10741-7

Gardner, J., Yu, R., Nguyen, Q., Brooks, C., & Kizilcec, R. (2023). Cross-institutional transfer learning for educational models: Implications for model performance, fairness, and equity. In Proceedings of the 2023 acm conference on fairness, accountability, and transparency (pp. 1664-1684). ACM. https://doi.org/10.1145/3593013.3594107

Griep, K., Stone, L., Edwards, S., Hill, M., Cui, J., & Liu, C. (2023). Ensuring ethical, transparent, and auditable use of education data and algorithms on AutoML. In Proceedings of the 2023 6th International Conference on Machine Learning and Natural Language Processing (pp. 66-72). https://doi.org/10.1145/3639479.3639492

Hoca, S., & Dimililer, N. (2025). A machine learning framework for student retention policy development: A case study. Applied Sciences, 15(6), 2989. https://doi.org/10.3390/app15062989

Hurshman, C. (2024). Do opaque algorithms have functions? Synthese, 204(3), 91. https://doi.org/10.1007/s11229-024-04745-2

Lopez-Muñoz, G. L., Serrano, C. G., & Sanchez-Ferreira, C. (2026). Student Dropout Prediction in Higher Education: A Systematic Review of Machine Learning Methods and Risk Factors. Journal of Information Technology Education: Research, 25, 20. https://doi.org/10.28945/5776

Khoudi, Z., Hafidi, N., Nachaoui, M., & Lyaqini, S. (2025). New approach to enhancing student performance prediction using machine learning techniques and clickstream data in virtual learning environments. SN Computer Science, 6(2), 139. https://doi.org/10.1007/s42979-024-03622-6

Kleimola, R., Hirsto, L., & Ruokamo, H. (2024). Promoting higher education students' self-regulated learning through learning analytics: A qualitative study. Education and Information Technologies, 30(4), 4959-4986. DOI: 10.1007/s10639-024-12978-4

Luo, D., Feng, C., Nong, Y., & Shen, Y. (2024). Autom3l: An automated multimodal machine learning framework with large language models. In Proceedings of the 32nd ACM International Conference on Multimedia (pp. 8586-8594). https://doi.org/10.1145/3664647.3680665

Milcu, T. (2025). Social and economic factors of academic dropout: A comprehensive analysis. Futurity of Social Sciences, 3(4), 341-353. https://doi.org/10.57125/FS.2025.12.20.14

Mnyawami, Y. N., Maziku, H. H., & Mushi, J. C. (2022). Comparative study of AutoML approach, conventional ensemble learning method, and K-Nearest Oracle-AutoML model for predicting student dropouts in Sub-Saharan African countries. Applied Artificial Intelligence, 36(1), 2145632. https://doi.org/10.1080/08839514.2022.2145632

Nigri, A., Bilancia, M., Cafarelli, B., & Magro, S. (2025). Modelling higher education dropouts using sparse and interpretable post-clustering logistic regression. arXiv preprint arXiv:2505.07582. https://doi.org/10.48550/arXiv.2505.07582

Permana, A. E., & Santoso, E. B. (2025). Student retention early warning procedures: A case study of higher education institution management. International Journal of Economics, Business and Accounting Research (IJEBAR), 9(3). https://doi.org/10.29040/ijebar.v9i3

Persad, D. (2025). Student Attrition Research [Dataset]. figshare. https://doi.org/10.6084/m9.figshare.28743122.v1

Rani, S. (2025). Studying the impact of anxiety, stress, and emotion on academic performance: A systematic review. Journal of Social, Humanity, and Education, 5(2), 131-141. doi:10.35912/jshe.v5i2.2437.

Reddy, T. S. (2025). Ethical dimensions of AutoML: Addressing bias, transparency, and responsible development. Journal of Computer Science and Technology Studies, 7(5), 981-988.

Rufibach, K. (2010). Use of Brier score to assess binary predictions. Journal of clinical epidemiology, 63(8), 938-939. https://doi.org/10.1016/j.jclinepi.2009.11.009

Schillizzi, C. (2025). Human-centered machine learning. In Handbook of Human-Centered Artificial Intelligence (pp. 1-38). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-97-8440-0_25-2

Schmidt, A., Cechinel, C., Queiroga, E. M., Primo, T., Ramos, V., & Bordin, A. S. (2025). Analyzing intervention strategies employed in response to automated academic-risk identification: A systematic review. IEEE Revista Iberoamericana de Tecnologias del Aprendizaje. DOI: 10.1109/RITA.2025.3540161

Schoina, C.-I., Gonida, S.-E., & Kiosseoglou, G. (2023). Dropout intentions in master studies: Applying Self-Determination Theory. Psychology: The Journal of the Hellenic Psychological Society, 28(1), 271–291.

Shivashankar, K., Hajj, G. S. A., & Martini, A. (2025). Scalability and maintainability challenges and solutions in machine learning: Systematic literature review. arXiv preprint arXiv:2504.11079. https://doi.org/10.48550/arXiv.2504.11079

Siemens, G., & Baker, R. S. D. (2012). Learning analytics and educational data mining: Towards communication and collaboration. In Proceedings of the 2nd international conference on learning analytics and knowledge (pp. 252-254). ACM. https://doi.org/10.1145/2330601.2330661

Swiss Cyber Institute. (2026). The EU AI Act: Implications for ethical AI in education. https://swisscyberinstitute.com/research/eu-ai-act-implications-ethical-ai-education/

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

Tinto, V. (2012). Leaving college: Rethinking the causes and cures of student attrition (2nd ed.). University of Chicago Press.

Ajjawi, R., Fischer, J., Adachi, C., Bearman, M., Boud, D., Macfarlane, S., … Coutts, K. (2026). Coming to belong in the university through small acts of connection. Studies in Higher Education, 51(1), 197–209. https://doi.org/10.1080/03075079.2025.2462666

Tsiakmaki, M., Kostopoulos, G., Kotsiantis, S., & Ragos, O. (2019). Implementing AutoML in educational data mining for prediction tasks. Applied Sciences, 10(1), 90. https://doi.org/10.3390/app10010090

Valencia-Arias, A., Valera Aredo, J. C., Valencia, J., Cardona-Acevedo, S., Patiño-Vanegas, J. C., & Uribe Bedoya, H. (2025). Key determinants of university dropout: Academic, economic, and psychosocial factors in student retention. In Frontiers in Education (Vol. 10, p. 1701644). Frontiers Media SA. https://doi.org/10.3389/feduc.2025.1701644

Wang, Y., Wang, H., Wang, S., Wind, S. A., & Gill, C. (2024). A systematic review and meta-analysis of self-determination-theory-based interventions in the education context. Learning and Motivation, 87, 102015. https://doi.org/10.1016/j.lmot.2024.102015

Weerts, H., Pfisterer, F., Feurer, M., Eggensperger, K., Bergman, E., Awad, N. (2024). Can fairness be automated? Guidelines and opportunities for fairness-aware AutoML. Journal of Artificial Intelligence Research, 79, 639-677. https://doi.org/10.1613/jair.1.14747