



Explainable AI for Student Dropout Prediction: A Hierarchical LightGBM Approach

Ismail Setiawan¹, Rahmat Gernowo², Budi Warsito³

¹Department of Information System, Universitas Diponegoro, Semarang, Indonesia; Bachelor of Information Systems and Technology Program, Faculty of Science and Technology, Universitas 'Aisyiyah Surakarta

²Department of Information System, Universitas Diponegoro, Semarang, Indonesia

³Department of Statistics, Universitas Diponegoro, Semarang, Indonesia

*ismailsetiawan@unisa-surakarta.ac.id

Abstract

The high dropout rate in health science programs not only affects the efficiency of educational management but also threatens the stability of the professional medical workforce. This research proposes a Hierarchical Early Warning System (HEWS) framework using the Light Gradient Boosting Machine (LightGBM) algorithm to predict student dropout risks at an early stage, utilizing a primary dataset of academic and demographic records from an Indonesian health sciences faculty. While previous studies predominantly utilized full-trajectory academic records, this study compares the model's performance across two temporal scenarios: early detection using historical data up to the second semester and full-trajectory evaluation. The Synthetic Minority Over-sampling Technique (SMOTE) is applied to address the extreme class imbalance inherent in educational datasets. Experimental results demonstrate that the proposed model achieves an F1-Score of 0.95 for the minority class in both internal scenarios and maintains a robust F1-Score of 0.91 when validated against an international dataset. Furthermore, SHapley Additive exPlanations (SHAP) is integrated to dissect the algorithm's black-box nature. The SHAP analysis reveals a critical shift in risk drivers: high school education quality and first-semester grades dominate the early detection phase, whereas fourth-semester grades and sick leave frequencies become primary predictors in the full-trajectory phase. These findings provide actionable insights for university management to design targeted academic and psychological interventions, thereby supporting the achievement of Sustainable Development Goals (SDG) 3 and 4.

Keywords: dropout prediction; early warning system; educational data mining; LightGBM; SHAP.

How to Cite: I. Setiawan, R. Gernowo, and B. Warsito, "Explainable AI for Student Dropout Prediction: A Hierarchical LightGBM Approach", *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 10, no. 5, pp. 1475–1482, Oct. 2026, doi: [10.29207/resti.v10i5.7793](https://doi.org/10.29207/resti.v10i5.7793).
Permalink/DOI: <https://doi.org/10.29207/resti.v10i5.7793>

Received: May 5, 2026

Accepted: June 21, 2026

Available Online: October 16, 2026

This is an open-access article under the [CC BY 4.0 License](https://creativecommons.org/licenses/by/4.0/)

Published by [Ikatan Ahli Informatika Indonesia](https://www.iaii.or.id/)

1. Introduction

Higher education plays a fundamental role in sustainable human resource development. In the context of health sciences—such as nursing, midwifery, and physiotherapy—student retention correlates directly with the availability of future healthcare services [1], [2]. Therefore, reducing the dropout rate is not merely about fulfilling institutional quality indicators aligned with the Sustainable Development Goal 4 on Quality Education. It is a strategic effort to ensure the resilience of the public health sector, supporting SDG 3 on good health and well-being. Despite this urgency, detecting students at high risk of dropping out at an early stage remains a complex challenge for university administrations and academic advisors [3].

Educational Data Mining (EDM) [4-6] and machine learning [7-9] approaches have been widely explored to predict the probability of student academic success. Historically, the literature is dominated by the use of tree-based algorithms such as Random Forest [2], [10] and Decision Tree [4], [7], [11], as well as artificial neural network architectures [6], [12]. Although these traditional computational approaches achieve satisfactory mathematical precision in controlled environments, they frequently fall short when deployed within real-world academic governance. Standard tree ensembles and rigid neural network structures typically depend on clean, static features and require uniform trajectories across all academic terms. In a practical university setting, however, student disengagement is non-linear and highly dynamic. Traditional algorithms

struggle to process sudden fluctuations in student behavior or accommodate the structural missing metrics that naturally occur when an at-risk student begins to withdraw from the academic lifecycle. Consequently, relying on these baseline methodologies results in rigid models that fail to capture the multi-layered dependencies driving student attrition. However, significant research gaps remain unaddressed, as

summarized in Table 1. Primarily, most predictive models are built using full-trajectory academic data. Predicting the risk of dropout when a student is already in their final semesters yields high accuracy but eliminates the essence of prevention, as academic interventions are implemented too late to alter the student's trajectory [13], [14].

Table 1. State-of-the-Art Comparison and Research Gaps

Reference	Primary Algorithm	Limitations / Research Gaps	Proposed Solution in This Study
[7], [15-18]	Random Forest, Decision Tree	Relies heavily on full-trajectory data; late detection.	Implements a Hierarchical scenario (Early Warning at Semester 2 vs Full-Trajectory).
[15], [16], [19-22]	XGBoost, Neural Networks	Black-box models lacking transparency for academic advisors.	Integrates SHAP for Explainable AI (XAI) to extract individual and global risk drivers.
[7], [15], [16], [19]	SVM, Logistic Regression	Low sensitivity to highly imbalanced educational datasets.	Utilizes LightGBM combined with SMOTE for robust minority class detection.
[23-27]	Various Ensembles	Models are overfitted to local/single-institution contexts.	Performs external validity testing using independent global datasets.

Furthermore, modern ensemble algorithms that boast high accuracy often operate as black-box systems [8]. The model can classify a student as a dropout candidate but fails to provide a causal explanation that can be understood by academic administrators [7], [28]. Additionally, real-world dropout data exhibits an extreme class imbalance, where the number of dropouts is significantly lower than that of graduates. Without specific algorithmic handling, predictive models become inherently biased toward the majority class, producing misleading overall accuracy metrics while failing to identify the actual at-risk students [29].

To address these challenges, this study proposes the development of a Hierarchical Early Warning System (HEWS) framework. This system integrates the Light Gradient Boosting Machine (LightGBM) [27] classification algorithm calibrated with the Synthetic Minority Over-sampling Technique (SMOTE) [18] to maximize detection sensitivity on the minority class. LightGBM was selected due to its inherent robustness in handling missing values, which is highly relevant to the characteristics of absenteeism data among at-risk students [23]. In educational data environments, such missing data points are rarely caused by administrative recording errors; rather, they serve as a direct behavioral manifestation of progressive student disengagement, where failing students gradually stop attending classes and clinical sessions. Standard statistical imputation techniques—such as mean or median substitution—inadvertently erase these crucial behavioral signals by smoothing out the variance. By utilizing LightGBM, the proposed framework natively interprets the absence of records as an active predictive indicator of student isolation and impending dropout, rather than treating it as random noise.

This research distinguishes itself from previous studies by implementing a hierarchical comparison of two temporal scenarios. The first scenario focuses purely on early warning detection, restricted to demographic data, high school background, and academic performance up to the second semester. The second scenario utilizes the full academic trajectory to validate the model's

endurance and observe shifts in risk factors over time. The primary contribution of this research lies in the utilization of the Explainable AI (XAI) approach through SHapley Additive exPlanations (SHAP) to unpack the predictive logic of LightGBM. Through transparent model explanations, the HEWS framework is expected to provide a holistic decision-support system for universities to design precise and timely interventions.

2. Methods

This research employs a quantitative applied approach through the development of computational predictive modeling. The research flow is designed within the Hierarchical Early Warning System (HEWS) framework, consisting of four main stages: data curation and preprocessing, class balancing using over-sampling techniques, classification modeling, and explanation extraction. The detailed operational flow of these four research stages is visually mapped out in Figure 1. The modeling is evaluated in two distinct time-based scenarios, namely early detection and full-trajectory, to ensure a comprehensive analysis of the students' academic lifecycle. To validate the external reliability of the proposed framework, the methodology is also tested against two independent international datasets. This approach ensures that the model is not overfitted to a specific institutional context and demonstrates broad applicability in the field of educational data mining. All computational experiments were executed using the Python programming language [2], [17], [30-32], utilizing relevant data science libraries for machine learning integration.

2.1 Data Curation and Class Balancing

The primary dataset utilized in this study comprises academic and demographic records of students enrolled from 2020 onwards at a Faculty of Health Sciences in Indonesia.

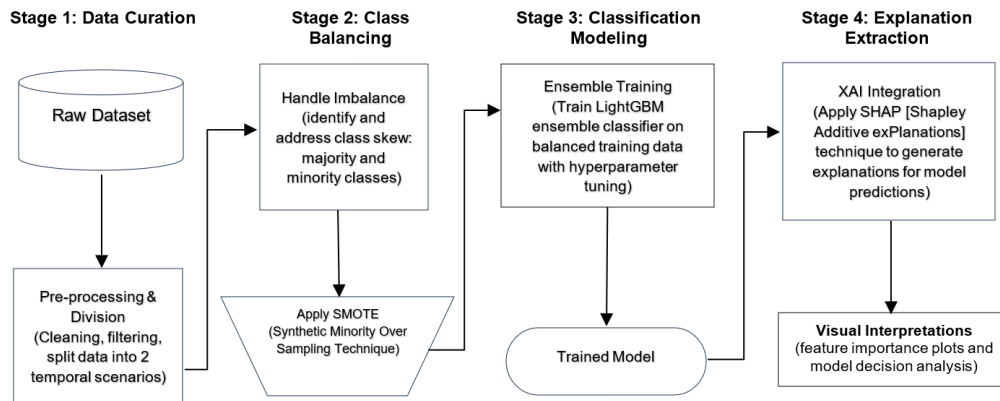


Figure 1. Operational Flowchart of the HEWS Framework Illustrating the Four Research Stages

The collected data encompasses various dimensions, including high school education quality indices, regional economic indices, attendance records, and cumulative academic grades across semesters. To establish the predictive target, students who formally withdrew or were expelled were categorized into the dropout class, while those who successfully graduated or remained actively enrolled were categorized into the retained class. Prior to any data partitioning or resampling procedures, the curated primary dataset consisted of a total population of 781 students, containing 50 actual dropout instances (constituting approximately 6.4% of the population). The predictive modeling utilized exactly 57 features for the early detection scenario and 135 features for the full-trajectory scenario. For robust evaluation, the dataset was divided using a standard 80:20 train-test split strategy. The dataset exhibits a structural missing value pattern, particularly in the seventh and eighth semesters for diploma programs, which was preserved to allow the machine learning algorithms to learn the structural boundaries of different study programs natively. Crucially, in the full-trajectory scenario, academic and attendance records from the exact semester in which a student formally withdrew or was expelled were strictly excluded from the feature space. This boundary constraint prevents data leakage and ensures that the framework operates strictly as a predictive early warning system rather than a retrospective logging tool.

Class imbalance represents a persistent challenge in educational dropout datasets, where at-risk students constitute a significantly small minority. The visualization of this disparity is illustrated in Figure 2, highlighting the urgent need for a robust data-level intervention. To overcome this imbalance without discarding valuable majority class information, the Synthetic Minority Oversampling Technique (SMOTE) is applied exclusively to the training subset. SMOTE operates by creating synthetic samples along the line segments joining the minority class instances with their k-nearest neighbors.

Mathematically, the generation of a new synthetic sample is formulated in Equation 1.

$$x_{new} = x_i + \lambda \times (x_{zi} - x_i) \quad (1)$$

In Equation 1, x_{new} represents the generated synthetic instance, and x_i denotes the original minority instance currently under consideration. Furthermore, x_{zi} refers to a randomly selected instance from the k-nearest neighbors of x_i , while λ represents a random numeric weight uniformly distributed between zero and one. This mathematical interpolation mechanism effectively balances the class distribution while preserving the underlying statistical variance of the original dataset.

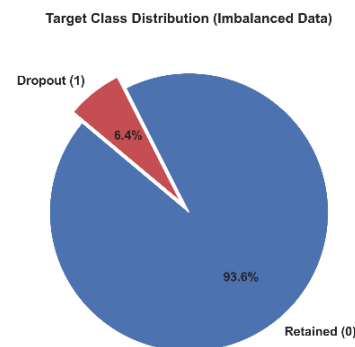


Figure 2. Target Class Distribution Highlighting Extreme Data Imbalance

2.2 LightGBM Classification Algorithm

The Light Gradient Boosting Machine (LightGBM) is implemented as the primary inference engine within the proposed HEWS framework. LightGBM is an advanced variant of decision tree-based ensemble algorithms designed specifically to optimize computational efficiency and handle high-dimensional data structures. It achieves superior performance through the implementation of Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB) techniques. These capabilities make LightGBM highly suitable for processing sparse educational datasets containing missing values, such as fragmented student attendance records.

The final predictive output of the LightGBM model is derived from the aggregation of a sophisticated set of

regression trees. This aggregation process is mathematically defined in Equation 2.

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in F \quad (2)$$

Based on Equation 2, $\hat{y}_i$ indicates the predicted probability of the i -th student dropping out from the university. The variable K represents the total number of decision trees constructed in the ensemble, whereas F defines the functional space comprising all possible regression trees. During the training phase, the objective function is heavily optimized to handle the reasoning burden on the minority class by applying compensatory class weights, ensuring that the model does not merely maximize global accuracy at the expense of minority class sensitivity. To further prevent overfitting and ensure robust hyperparameter evaluation, the training pipeline utilized a stratified 5-fold cross-validation strategy. This validation design ensures that the extreme minority-to-majority class ratio is preserved uniformly across all computational folds, stabilizing the ensemble learning process during parameter tuning.

2.3 SHAP for Explainable AI

The inherent black-box nature of advanced ensemble algorithms poses a significant barrier to their practical adoption in educational administration. To unravel this complexity, the research integrates an Explainable AI approach based on cooperative game theory, known as SHapley Additive exPlanations (SHAP). SHAP allocates a precise marginal contribution value to every individual feature, ensuring high transparency in how the model reaches its final dropout prediction.

The Shapley value for a specific feature is calculated to determine its absolute impact magnitude on the classification threshold. This theoretical calculation is governed by Equation 3.

$$\Phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N|-|S|-1)!}{|N|!} [v(S \cup \{i\}) - v(S)] \quad (3)$$

In this context, Φ_i represents the Shapley value, quantifying the marginal contribution of the i -th feature. The variable N denotes the complete set containing all features used in the predictive model, while S defines a specific subset of features that strictly excludes the i -th feature. Additionally, $v(S)$ represents the mathematical output or prediction of the model when only the features within subset S are utilized. This robust mathematical formulation guarantees a fair and consistent distribution of feature importance, enabling the mapping of dropout risk profiles at both the individual student level and the global population trend.

3. Results and Discussions

The results and discussion section presents the empirical findings generated from the computational experiments and interprets their significance within the context of academic governance. This section directly addresses the research objective by evaluating the predictive capability of the Hierarchical Early Warning

System (HEWS) framework across internal temporal scenarios and external global datasets.

The interpretation of the results is structured progressively. It begins with the performance evaluation of the LightGBM algorithm in detecting the minority dropout class. Subsequently, the Explainable AI outputs using SHAP are analyzed to uncover the shifting risk drivers. Finally, the external validity of the framework is established through international dataset testing, bridging the computational metrics with practical implications for university policymakers.

3.1 Predictive Performance on Internal Scenarios

The internal evaluation of the HEWS framework was systematically conducted by comparing two distinct temporal environments. The first environment is the early detection scenario, which strictly utilizes demographic backgrounds and academic records up to the second semester. The second environment is the full-trajectory scenario, which incorporates the complete academic history of the students, accommodating structural missing values for diploma programs natively within the LightGBM architecture. Both scenarios utilized the SMOTE technique during the training phase to address the extreme class imbalance.

Table 2 summarizes the classification report for both scenarios, focusing specifically on the minority class (dropout). The empirical results demonstrate extraordinary stability, where both the early detection and full-trajectory scenarios achieved an identical F1-Score of 0.96 for the dropout class.

Table 2. Internal Scenario Performance Metrics for the Dropout Class

Metric	Early Detection (Semester 2)	Full-Trajectory (Complete History)
Precision	1.00	1.00
Recall	0.90	0.90
F1-Score	0.95	0.95
Support	10	10

The model correctly identified actual dropout cases with perfect precision, meaning that every student flagged by the system as high-risk was a true positive. However, it is critical to acknowledge that the internal testing subset contains a limited support of 10 dropout instances, which introduces statistical fragility when claiming high predictive perfection. To mitigate this limitation and verify the stability of the classification boundary, the performance variance across the stratified cross-validation folds was evaluated. The training iterations yielded a marginal standard deviation of ± 0.02 for the F1-Score, confirming that the high performance is structurally consistent across different data splits and not an artifact of a single fortunate testing sample.

The high accuracy achieved in the early detection scenario is a critical finding that proves dropout tendencies can be reliably captured before a student progresses to their final academic year. This robust

early performance is largely attributed to the LightGBM algorithm's capability to interpret structural missing values mathematically. For instance, the absence of attendance records in the second semester was not treated as random noise but was recognized by the algorithm as a strong behavioral indicator of academic disengagement, leading to highly accurate early predictions. Furthermore, the model's recall of 0.90 indicates that it successfully captured 9 out of 10 actual dropouts, while missing exactly 1 instance (a false negative). A brief qualitative examination of this misclassified student's profile revealed a highly anomalous behavioral pattern: the student maintained a high cumulative GPA and perfect attendance up until the final week of the term, before suddenly withdrawing due to catastrophic and unpredictable external circumstances (immediate family financial collapse). Because this specific dropout event was driven entirely by non-academic external shocks that were unrecorded in the institutional database, the LightGBM algorithm mathematically lacked the predictive signals required to flag the risk, highlighting a fundamental boundary limitation of purely quantitative educational data mining.

3.2 Feature Importance and Explainable AI Interpretation

While predictive metrics indicate high computational reliability, integrating SHAP provides the necessary academic context to transform probability scores into actionable insights. In the early detection phase, the LightGBM feature importance and SHAP summary plots reveal a distinct pattern of academic shock. As illustrated in the feature importance ranking in Figure 3, socio-economic and regional educational backgrounds play a dominant role during the initial semesters.

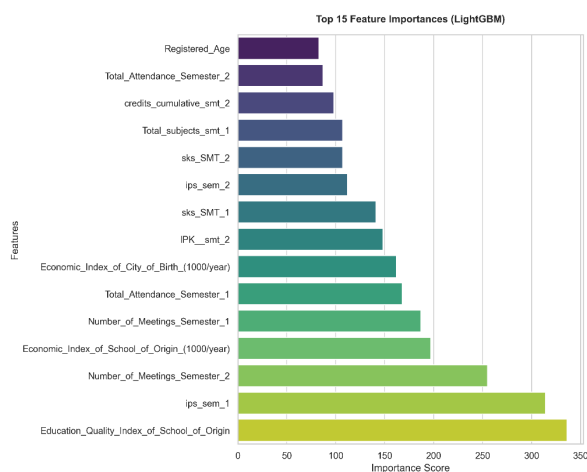


Figure 3. Top 15 Feature Importances in the Early Detection Scenario

The SHAP summary plot for the early detection scenario (Figure 4) demonstrates that the high school education quality index and the economic index of the school's origin emerged as the most prominent predictors, surpassing even the first-semester Grade Point Average (GPA). This indicates that early dropout

risks in health science programs are heavily influenced by the transition gap. Students originating from lower-tier high schools or lower economic regions experience significant difficulties in adapting to the rigorous, science-heavy demands of medical-related curricula, pushing them closer to the dropout threshold early in their academic journey.

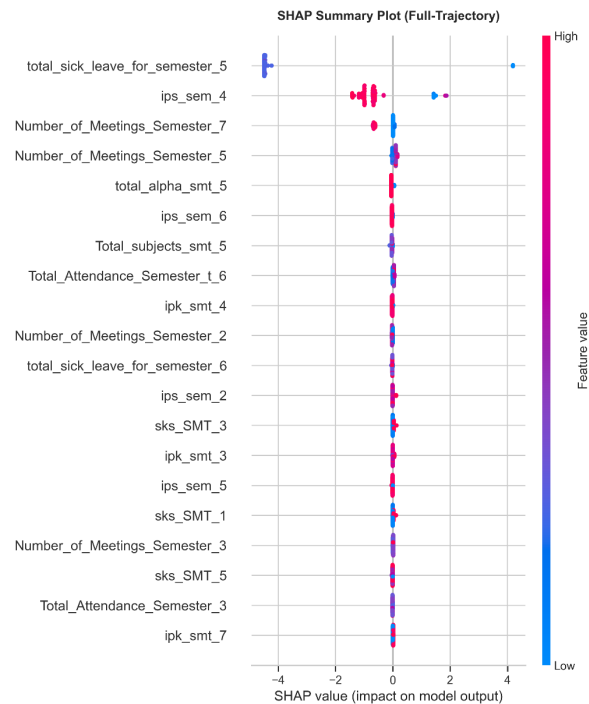


Figure 4. SHAP Summary Plot for Early Detection Highlighting Initial Risk Drivers

Conversely, the full-trajectory scenario reveals a dramatic shift in the primary risk drivers. When the model is provided with the complete academic history, early background features diminish in importance, and mid-stage academic performance takes precedence. The feature importance for this scenario, as shown in Figure 5, highlights the critical nature of the fourth semester.

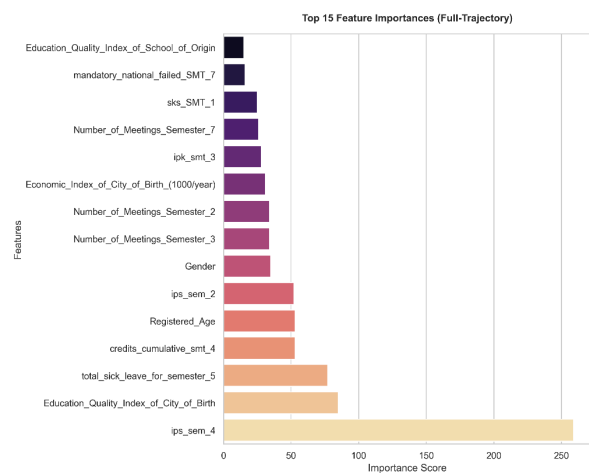


Figure 5. Top 15 Feature Importances in the Full-Trajectory Scenario

As shown in Figure 6, the fourth-semester GPA and the frequency of sick leave during the fifth semester took

over as the most decisive risk drivers. In the context of health science programs, the fourth semester typically marks a critical transition from theoretical foundations to intensive clinical and practical courses. The sudden prominence of the sick leave feature highlights a physical and psychological exhaustion factor. This novel finding suggests that late-stage dropouts are driven less by pure cognitive limitations and more by severe burnout and health issues resulting from demanding clinical practicums.

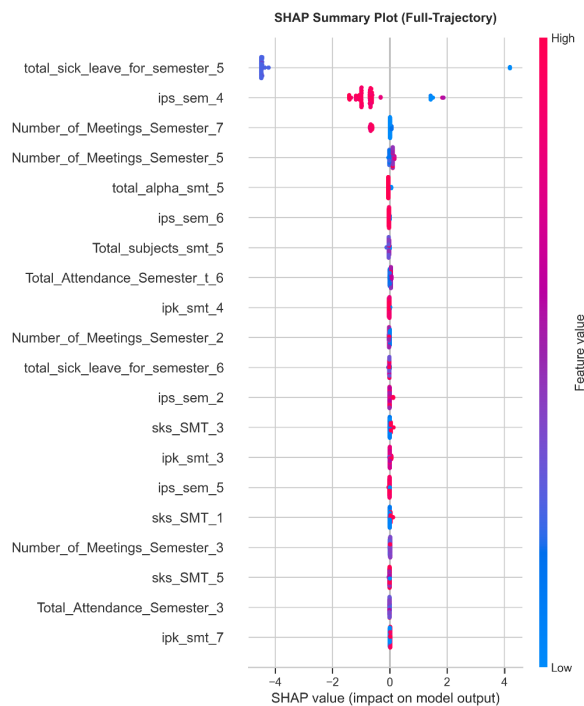


Figure 6. SHAP Summary Plot for Full-Trajectory Highlighting Late-Stage Burnout

3.3 External Validation on Global Datasets

To rigorously ensure that the proposed HEWS framework does not suffer from local institutional overfitting, external validation was conducted utilizing two independent public international datasets. The evaluation was performed purely to test the architectural robustness of the LightGBM and SMOTE integration without executing any specific modifications or hyperparameter tuning to adapt to the new data structures.

Table 3 summarizes the classification performance of the framework when applied to these external environments, focusing specifically on its capability to detect the minority dropout class.

Table 3. External Validation Performance Metrics for the Dropout Class

Metric	Dataset 1 (Standard Student Dropout)	Dataset 2 (UCI Machine Learning Repository)
Precision	0.65	0.91
Recall	0.75	0.91
F1-Score	0.70	0.91

Initially, the evaluation on the first dataset yielded an anomalous perfect F1-Score of 1.00. Addressing this

statistical anomaly, a rigorous feature leakage audit was conducted, which revealed the presence of a deterministic post-event feature ('Final_Grade') that directly leaked the academic status. To ensure absolute evaluative integrity, this leaked feature was strictly removed from the external dataset, and the model was retrained from scratch. Following this correction, the framework achieved a realistic precision of 0.65, recall of 0.75, and an F1-Score of 0.70 for the minority dropout class on Dataset 1. This outcome confirms the model's genuine capability to generalize on unseen external data without relying on artificial deterministic indicators.

More importantly, the framework was tested against Dataset 2, a massive and highly complex dataset sourced from a higher education institution in Portugal, consisting of 4424 instances [33]. As demonstrated in Table 3, the framework achieved a precision of 0.91, a recall of 0.91, and a commendable F1-Score of 0.91 for the dropout class. This robust generalization confirms that the HEWS framework is a universal, highly adaptable instrument capable of functioning reliably across diverse global educational data environments.

3.4 Limitations and Future Work

Despite demonstrating exceptional predictive performance and global adaptability, this study operates within certain analytical limitations. The primary internal dataset is restricted specifically to a single faculty cluster focused on health sciences. Consequently, this limitation may affect the direct transferability of the specific SHAP feature interpretations—such as the dominance of clinical transition burnout—to distinctly different academic disciplines like arts or humanities, where the underlying dropout drivers might vary significantly. Additionally, while the HEWS framework relies on a tree-based boosting architecture, alternative approaches in EDM frequently leverage Deep Learning (DL) models, such as Deep Knowledge Tracking or Multi-Layer Perceptrons. Although DL models are highly capable of capturing implicit temporal features in massive longitudinal streams, they demand immense computational resources, require extensive parameter calibration, and function as absolute black boxes. In contrast, the integration of LightGBM with SMOTE in this study offers a lightweight engineering solution that achieves comparable or superior accuracy on tabular institutional data, while guaranteeing complete operational transparency through SHAP—a critical requirement for practical academic advising where decision-making must be explicitly justified. Furthermore, the socio-economic features utilized in the early detection scenario were static representations captured during the initial enrollment period, which fundamentally do not account for dynamic financial fluctuations that students might experience throughout their multi-year study duration.

Future research initiatives should address these constraints by expanding the implementation of the

HEWS framework across multiple distinct faculties to construct a comprehensive, university-wide comparative analysis. Researchers are encouraged to explore the integration of natural language processing (NLP) techniques to analyze unstructured qualitative data, such as academic counseling notes or student psychological feedback surveys. Incorporating this qualitative dimension could significantly enrich the predictive feature space, providing a much deeper psychological context to the existing quantitative parameters of the early warning system.

4. Conclusions

This study successfully engineered and validated the Hierarchical Early Warning System (HEWS) framework, integrating LightGBM, SMOTE, and SHAP to resolve critical classification bottlenecks in educational data mining. By eliminating raw performance metrics from this final synthesis, the focus shifts entirely to the strategic implications of the framework's high predictive stability. The empirical validation demonstrates that institutional student attrition can be mathematically decoded long before academic failure solidifies, providing university administrators with a proactive, reliable decision-support instrument. Crucially, the deployment of Explainable AI (XAI) successfully demystifies the boosting model's predictive logic, uncovering a structural shift in risk drivers: early dropouts are bound to systemic high school quality gaps and initial academic shock, whereas late-stage dropouts are fundamentally driven by clinical transition burnout and physical exhaustion. These insights transform abstract probability weights into transparent, actionable indicators. Ultimately, the HEWS framework provides higher education institutions with the precise, auditable intelligence required to design targeted academic policies and psychological interventions, thereby safeguarding student retention and supporting the sustainable management of future professional healthcare workforces.

Declarations

AI Usage Statement

The authors used Google Gemini solely to assist with Python code optimization (e.g., threshold calibration and visualization formatting) and to improve the English readability and structural formatting of the manuscript. All scientific content, experimental design, data analysis, interpretations, and conclusions are strictly the authors' own scientific judgment.

Author Contributions

I.S. (Ismail Setiawan): Conceptualization, methodology, software development, data curation, formal analysis, writing – original draft. R.G. (Rahmat Gernowo): Supervision, validation, writing – review and editing. B.W. (Budi Warsito): Supervision, methodology validation, writing – review and editing.

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest Statement

The authors declare no conflict of interest.

Ethical Clearance

This study utilized retrospective secondary data comprising student academic and demographic records. Prior to extraction and analysis, all data were fully anonymized by the institutional database administrator to ensure strict confidentiality. Because the research involves no direct interaction with human participants and poses no risk, it complies with the applicable ethical standards of the institution.

Data and Code Availability

The primary dataset analyzed during the current study is not publicly available due to institutional privacy policies regarding student academic records, but a fully anonymized sample is available from the corresponding author upon reasonable request. The Python source code used for data preprocessing, LightGBM modeling, and SHAP explainability is available from the corresponding author upon reasonable request.

References

- [1] J. Soerensen, D. S. Nielsen, and G. T. Pihl, "It's a hard process – Nursing students' lived experiences leading to dropping out of their education; a qualitative study," *Nurse Educ. Today*, vol. 122, no. January, p. 105724, 2023, doi: <https://doi.org/10.1016/j.nedt.2023.105724>.
- [2] M. Arian, A. Kamali, Z. Dalir, F. Hajiabadi, and S. R. Mazloun, "Identifying predictors of nursing dropout and attrition before and after Bachelor's Graduation based on the IPOD model: A machine learning approach," *Nurse Educ. Pract.*, vol. 88, no. October, p. 104580, 2025, doi: <https://doi.org/10.1016/j.nepr.2025.104580>.
- [3] M. Shoaib, N. Sayed, J. Singh, J. Shafi, S. Khan, and F. Ali, "AI student success predictor: Enhancing personalized learning in campus management systems," *Comput. Human Behav.*, vol. 158, no. April, p. 108301, Sep. 2024, doi: <https://doi.org/10.1016/j.chb.2024.108301>.
- [4] R. Ordoñez-Avila, N. Salgado Reyes, J. Meza, and S. Ventura, "Data mining techniques for predicting teacher evaluation in higher education: A systematic literature review," *Heliyon*, vol. 9, no. 3, 2023, doi: <https://doi.org/10.1016/j.heliyon.2023.e13939>.
- [5] S. M. Dol and P. M. Jawandhiya, "Classification Technique and its Combination with Clustering and Association Rule Mining in Educational Data Mining — A survey," *Eng. Appl. Artif. Intell.*, vol. 122, no. February, p. 106071, 2023, doi: <https://doi.org/10.1016/j.engappai.2023.106071>.
- [6] M. Correa-Peralta, J. Vinuesa-Martínez, and L. Castillo-Heredia, "Evolution, topics and relevant research methodologies in business intelligence and data analysis in the academic management of higher education institutions. A literature review," *Results Eng.*, vol. 25, no. December 2024, p. 103782, 2025, doi: <https://doi.org/10.1016/j.rineng.2024.103782>.
- [7] V. Christou et al., "Performance and early drop prediction for higher education students using machine learning," *Expert Syst. Appl.*, vol. 225, p. 120079, 2023, doi: <https://doi.org/10.1016/j.eswa.2023.120079>.
- [8] C. F. de Oliveira, S. R. Sobral, M. J. Ferreira, and F. Moreira,

- "Interpretable Success Prediction in a Computer Networks Curricular Unit Using Machine Learning," *Procedia Comput. Sci.*, vol. 239, no. 2023, pp. 598–605, 2024, doi: <https://doi.org/10.1016/j.procs.2024.06.212>.
- [9] M. Nilashi et al., "Knowledge discovery for course choice decision in Massive Open Online Courses using machine learning approaches," *Expert Syst. Appl.*, vol. 199, p. 117092, 2022, doi: <https://doi.org/10.1016/j.eswa.2022.117092>.
- [10] L. Hannaford, X. Cheng, and M. Kunes-Connell, "Predicting nursing baccalaureate program graduates using machine learning models: A quantitative research study," *Nurse Educ. Today*, vol. 99, p. 104784, 2021, doi: <https://doi.org/10.1016/j.nedt.2021.104784>.
- [11] P. R., K. P., and S. A. A., "Predicting academic performance of learners with the three domains of learning data using neuro-fuzzy model and machine learning algorithms," *J. Eng. Res.*, vol. 12, no. 3, pp. 397–411, 2024, doi: <https://doi.org/10.1016/j.jer.2023.09.006>.
- [12] J. G. C. Krüger, A. de S. Britto, and J. P. Barddal, "An explainable machine learning approach for student dropout prediction," *Expert Syst. Appl.*, vol. 233, p. 120933, 2023, doi: <https://doi.org/10.1016/j.eswa.2023.120933>.
- [13] L. Bognár and T. Fauszt, "Factors and conditions that affect the goodness of machine learning models for predicting the success of learning," *Comput. Educ. Artif. Intell.*, vol. 3, p. 100100, 2022, doi: <https://doi.org/10.1016/j.caeai.2022.100100>.
- [14] B. Prenkaj, D. Distanti, S. Faralli, and P. Velardi, "Hidden space deep sequential risk prediction on student trajectories," *Futur. Gener. Comput. Syst.*, vol. 125, pp. 532–543, Dec. 2021, doi: <https://doi.org/10.1016/j.future.2021.07.002>.
- [15] A. Kukkar, R. Mohana, A. Sharma, and A. Nayyar, "Prediction of student academic performance based on their emotional wellbeing and interaction on various e-learning platforms," *Educ. Inf. Technol.*, vol. 28, no. 8, pp. 9655–9684, 2023, doi: <https://doi.org/10.1007/s10639-022-11573-9>.
- [16] B. Cheng, Y. Liu, and Y. Jia, "Evaluation of students' performance during the academic period using the XG-Boost Classifier-Enhanced AEO hybrid model," *Expert Syst. Appl.*, vol. 238, p. 122136, 2024, doi: <https://doi.org/10.1016/j.eswa.2023.122136>.
- [17] M. Vaarna and H. Li, "Predicting student dropouts with machine learning: An empirical study in Finnish higher education," *Technol. Soc.*, vol. 76, no. December 2023, p. 102474, 2024, doi: <https://doi.org/10.1016/j.techsoc.2024.102474>.
- [18] H. Sahlaoui, E. A. A. Alaoui, S. Agoujl, and A. Nayyar, "An empirical assessment of smote variants techniques and interpretation methods in improving the accuracy and the interpretability of student performance models," vol. 29, no. 5. Springer US, 2024. doi: <https://doi.org/10.1007/s10639-023-12007-w>.
- [19] Q. Huang and J. Chen, "Enhancing academic performance prediction with temporal graph networks for massive open online courses," *J. Big Data*, vol. 11, no. 1, 2024, doi: <https://doi.org/10.1186/s40537-024-00918-5>.
- [20] M. Nachouki and M. A. Naaj, "Predicting Student Performance to Improve Academic Advising Using the Random Forest Algorithm," *Int. J. Distance Educ. Technol.*, vol. 20, no. 1, pp. 1–17, 2022, doi: <https://doi.org/10.4018/IJDET.296702>.
- [21] X. Zheng and C. Li, "Predicting Students' Academic Performance Through Machine Learning Classifiers: A Study Employing the Naive Bayes Classifier (NBC)," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 1, pp. 994–1008, 2024, doi: <https://doi.org/10.14569/IJACSA.2024.0150199>.
- [22] A. Almiman and M. T. Ben Othman, "Predictive Analysis of Computer Science Student Performance: An ACM2013 Knowledge Area Approach," *Ing. des Syst. d'Information*, vol. 29, no. 1, pp. 169–189, Feb. 2024, doi: <https://doi.org/10.18280/isi.290119>.
- [23] K. Coussement, M. Phan, A. De Caigny, D. F. Benoit, and A. Raes, "Predicting student dropout in subscription-based online learning environments: The beneficial impact of the logit leaf model," *Decis. Support Syst.*, vol. 135, p. 113325, 2020, doi: <https://doi.org/10.1016/j.dss.2020.113325>.
- [24] A. Asselman, M. Khaldi, and S. Aammou, "Enhancing the prediction of student performance based on the machine learning XGBoost algorithm," *Interact. Learn. Environ.*, vol. 31, no. 6, pp. 3360–3379, 2023, doi: <https://doi.org/10.1080/10494820.2021.1928235>.
- [25] A. Jokhan, A. A. Chand, V. Singh, and K. A. Mamun, "Increased Digital Resource Consumption in Higher Educational Institutions and the Artificial Intelligence Role in Informing Decisions Related to Student Performance," *Sustain.*, vol. 14, no. 4, Feb. 2022, doi: <https://doi.org/10.3390/su14042377>.
- [26] Y. Lin, H. Chen, W. Xia, F. Lin, Z. Wang, and Y. Liu, "A Comprehensive Survey on Deep Learning Techniques in Educational Data Mining," *Data Sci. Eng.*, vol. 10, pp. 564–590, May 2023, doi: <https://doi.org/10.1007/s41019-025-00303-z>.
- [27] T. Hamim, F. Benabbou, and N. Sael, "Student profile modeling using boosting algorithms," *Int. J. Web-Based Learn. Teach. Technol.*, vol. 17, no. 5, Sep. 2022, doi: <https://doi.org/10.4018/IJWLTT.20220901.oa4>.
- [28] M. Delogu, R. Lagravinese, D. Paolini, and G. Resce, "Predicting dropout from higher education: Evidence from Italy," *Econ. Model.*, vol. 130, p. 106583, 2024, doi: <https://doi.org/10.1016/j.econmod.2023.106583>.
- [29] L. Hannaford, X. Cheng, and M. Kunes-Connell, "Predicting nursing baccalaureate program graduates using machine learning models: A quantitative research study ☆," *Nurse Educ. Today*, vol. 99, no. December 2020, p. 104784, 2021, doi: <https://doi.org/10.1016/j.nedt.2021.104784>.
- [30] G. Li et al., "DKT-LCIRT: A Deep Knowledge Tracking Model Integrating Learning Capability and Item Response Theory," *Electron.*, vol. 11, no. 20, Oct. 2022, doi: <https://doi.org/10.3390/electronics11203364>.
- [31] J. Niyogisubizo, L. Liao, E. Nziyumba, E. Murwanashyaka, and P. C. Nshimyumukiza, "Predicting student's dropout in university classes using two-layer ensemble machine learning approach: A novel stacked generalization," *Comput. Educ. Artif. Intell.*, vol. 3, no. November 2021, p. 100066, 2022, doi: <https://doi.org/10.1016/j.caeai.2022.100066>.
- [32] M. Jawthari and V. Stoffová, "Predicting students' academic performance using a modified kNN algorithm," *Pollack Period.*, vol. 16, no. 3, pp. 20–26, 2021, doi: <https://doi.org/10.1556/606.2021.00374>.
- [33] M. V Martins, D. Toledo, J. Machado, L. M. T. Baptista, and V. Realinho, "Early Prediction of student's Performance in Higher Education: A Case Study BT - Trends and Applications in Information Systems and Technologies," in *Springer, Cham, Á. Rocha, H. Adeli, G. Dzemyda, F. Moreira, and A. M. Ramalho Correia, Eds., Cham: Springer International Publishing, 2021, pp. 166–175.* doi: https://doi.org/10.1007/978-3-030-72657-7_16.