



Using Explainable Artificial Intelligence to Predict Student Dropout Risk from Learning Management System Behavioural Data Early

Bode Bawasah

University of Waterloo, Ontario, Canada

DOI : <https://doi.org/10.55248/gengpi.07.0926.2922>

ABSTRACT :

Student attrition remains a persistent challenge in higher education, affecting institutional performance, educational equity, resource allocation, and learners' long-term academic outcomes. The expansion of digital learning environments has created extensive behavioural data that can reveal changes in student engagement before conventional academic indicators identify withdrawal risk. This study proposes an explainable artificial intelligence framework for early prediction of student dropout using longitudinal behavioural data captured through Learning Management Systems. The framework integrates login frequency, session regularity, content access, assessment submissions, discussion participation, inactivity duration, learning-resource consumption, and temporal engagement changes to construct dynamic student-risk profiles. Machine-learning models are developed to estimate dropout probability at successive stages of the academic period, while explainability techniques identify the behavioural factors contributing to individual predictions. Particular emphasis is placed on distinguishing temporary disengagement from persistent behavioural deterioration and determining how early reliable warnings can be generated. Model performance is evaluated using discrimination, calibration, class-imbalance-sensitive metrics, prediction lead time, and explanation consistency. The proposed approach supports transparent early-warning systems that enable institutions to identify emerging disengagement patterns and implement timely, evidence-informed student support interventions.

Keywords: Explainable Artificial Intelligence; Student Dropout Prediction; Learning Management Systems; Behavioural Analytics; Early Warning Systems; Student Engagement

1. INTRODUCTION

1.1 Student Attrition in Digitally Mediated Higher Education

Student attrition remains a persistent challenge for higher education because withdrawal disrupts academic progression, reduces institutional retention, creates financial consequences, and can intensify inequalities among learners with different socioeconomic and educational circumstances [1]. Traditional approaches to identifying vulnerable students frequently depend on assessment outcomes, attendance records, or formal withdrawal indicators that become available after meaningful disengagement has already occurred [2]. The increasing integration of Learning Management Systems (LMSs) into university teaching has created a richer environment for observing student engagement continuously [3]. Every login, resource interaction, assessment attempt, submission, and online learning activity can generate timestamped behavioural evidence [5]. Consequently, LMS data provide an opportunity to move student-retention analytics from retrospective identification toward continuous recognition of behavioural changes that may precede academic withdrawal [4].

1.2 Behavioural Signals Preceding Student Dropout

Student disengagement is rarely represented by a single digital action. Instead, emerging risk may become visible through declining login frequency, longer inactivity periods, increasingly irregular study sessions, reduced learning-content access, missed submissions, and decreasing participation in online discussions [5]. Changes in resource-consumption patterns may additionally reveal weakening academic engagement before substantial deterioration appears in formal grades [6]. Importantly, these behaviours acquire greater predictive meaning when examined longitudinally rather than as isolated activity counts. A brief period of inactivity may reflect temporary circumstances, whereas persistent or accelerating disengagement across successive learning periods can indicate a developing withdrawal trajectory [7]. Dropout prediction should therefore capture the direction, persistence, recurrence, and magnitude of behavioural change over time, rather than treating engagement as static.

1.3 Limitations of Conventional and Black-Box Prediction Approaches

Conventional student-risk identification often relies on demographic characteristics, historical academic performance, cumulative grades, or predefined thresholds for attendance and assessment completion [4]. Although informative, these indicators may inadequately represent rapidly changing engagement within digital learning environments. Machine-learning techniques can detect more complex relationships among behavioural variables and potentially recognize risk earlier [8]. However, predictive improvement introduces another challenge when sophisticated models operate as black boxes. A high dropout probability alone does not explain which behaviours generated the prediction or what support may be appropriate. This opacity can reduce educator confidence, complicate intervention decisions, and make potentially consequential predictions difficult to justify [9]. Effective early-warning systems therefore require both predictive capability and transparent explanations connecting risk estimates to observable student behaviours.

1.4 Research Aim, Objectives, and Contribution

This study develops an explainable artificial intelligence framework for predicting student dropout risk early using longitudinal LMS behavioural data. Rather than asking only which students are likely to withdraw, the framework addresses two additional questions: why is a particular student classified as at risk, and how early can a reliable warning be generated? Behavioural trajectories are constructed from temporally ordered LMS interactions and evaluated using predictive models capable of estimating evolving dropout probabilities [5]. Explainability mechanisms subsequently identify the behavioural factors contributing to individual risk estimates [9]. The study contributes an integrated approach combining temporal engagement modelling, early-warning evaluation, student-level explanation, and intervention-oriented risk interpretation, thereby positioning predictive analytics as transparent decision support for timely, evidence-informed student retention strategies.

2. LITERATURE REVIEW AND CONCEPTUAL FOUNDATIONS

2.1 Student Engagement, Disengagement, and Dropout Dynamics

Student engagement is multidimensional, encompassing behavioural participation, academic involvement, and interaction with instructors, peers, assessments, and learning resources [7]. Within digitally mediated education, behavioural engagement becomes observable through the frequency, consistency, and sequencing of activities performed within an LMS [9]. Academic engagement is reflected in assessment attempts, submission behaviour, and sustained interaction with course materials, while interaction-based engagement includes participation in forums and other collaborative learning activities [8]. Disengagement can consequently emerge as a gradual deterioration across several dimensions rather than an immediate decision to withdraw. Declining participation, increasing inactivity, and disrupted study routines may accumulate across successive weeks before formal dropout becomes evident [11]. End-of-semester grades therefore provide comparatively late evidence of difficulties that may have developed considerably earlier. Longitudinal behavioural trajectories offer a stronger temporal perspective by revealing whether engagement is stable, recovering, intermittently disrupted, or progressively deteriorating before withdrawal [10].

2.2 LMS Behavioural Analytics for Student-Risk Identification

LMS platforms generate detailed digital traces that permit engagement to be examined at substantially greater temporal resolution than conventional academic records [12]. Login frequency and recency can indicate continuity of platform engagement, while session duration and regularity characterize the consistency of study behaviour. Content views, resource downloads, navigation sequences, and repeated access to instructional materials provide additional evidence about learning-resource consumption [9]. Assignment submission timing and assessment participation reveal task-oriented engagement, whereas forum interactions capture elements of collaborative and communicative participation [13]. Prolonged inactivity intervals may be particularly informative when they deviate from a student's established learning routine. Nevertheless, isolated activity counts can produce misleading interpretations. Ten logins in one week, for example, cannot establish whether engagement is improving or deteriorating without historical context [11]. Longitudinal analytics instead examine changes between successive observation periods, allowing persistent decline, behavioural volatility, recovery, recurrence, and deviations from individual engagement baselines to become measurable risk signals [14].

2.3 Machine Learning for Student Dropout Prediction

Machine-learning methods provide different levels of complexity for transforming behavioural indicators into dropout-risk estimates. Logistic regression offers an interpretable statistical baseline, while decision trees generate understandable decision structures but can become unstable with complex datasets [8]. Random forests improve robustness through ensemble learning, whereas gradient-boosting models can capture nonlinear interactions among numerous behavioural predictors [12]. Support vector machines are effective for complex classification boundaries, and neural networks can learn higher-order representations from large behavioural datasets [15]. However, predictive modelling presents methodological challenges. Dropout cases may constitute a minority class, making accuracy alone potentially misleading and requiring greater attention to recall, precision, and class-sensitive evaluation [10]. Model performance can also deteriorate across courses, cohorts, or institutions when behavioural distributions differ. Furthermore, many studies emphasize predictive performance without evaluating when a prediction becomes reliable, whether educators can understand it, or whether sufficient time remains for meaningful intervention [13].

2.4 Explainable AI and the Research Gap

Explainable artificial intelligence provides mechanisms for examining how predictive models transform student behaviour into estimated dropout risk [14]. Global explainability identifies variables that influence predictions across the overall student population, providing insight into whether inactivity, submission delays, resource engagement, or other behavioural characteristics dominate model decisions. Local explainability addresses a different question by identifying the factors responsible for an individual student's prediction [9]. Feature-importance methods can rank influential predictors, while SHapley Additive exPlanations (SHAP) can quantify the direction and relative contribution of individual features to a particular risk estimate [15]. Such explanations are especially important in educational settings because identical dropout probabilities may arise from substantially different behavioural trajectories. One student may exhibit persistent inactivity, whereas another may remain active but repeatedly miss assessments [11]. Existing predictive research therefore leaves an important gap: high classification performance does not establish why risk is increasing, whether the explanation is educationally meaningful, or whether the warning occurs early enough to permit intervention [13]. Integrating temporal prediction with student-level explainability is consequently necessary for converting algorithmic risk scores into interpretable early-warning intelligence.

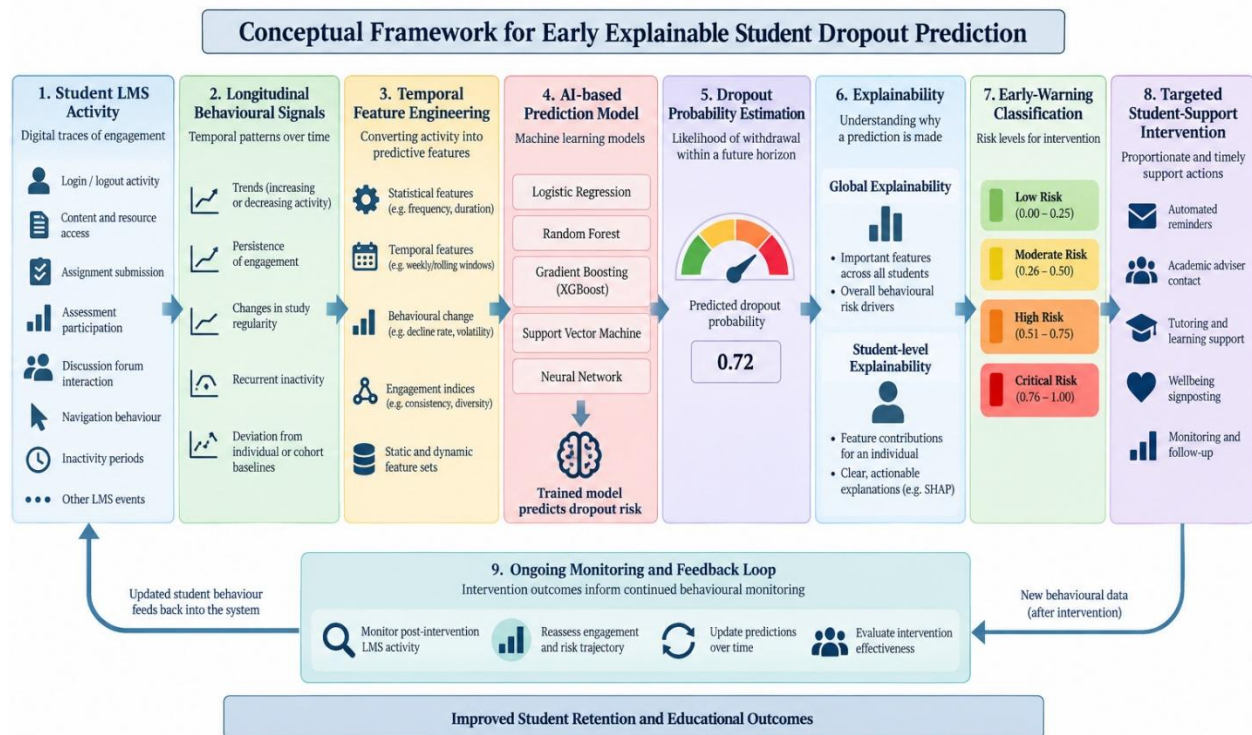


Figure 1. Conceptual Framework for Early Explainable Student Dropout Prediction

3. LMS DATA ACQUISITION, TEMPORAL INTEGRATION, AND BEHAVIOURAL FEATURE ENGINEERING

3.1 Study Population and LMS Data Acquisition

The study population comprises students enrolled in digitally supported undergraduate courses in which the LMS functions as a primary environment for distributing learning materials, submitting assessments, communicating, and recording academic interactions [14]. Observation begins from the first instructional week and continues until course completion or confirmed withdrawal, enabling behavioural development to be reconstructed chronologically [16]. Eligible records require active course enrolment, a valid anonymized student identifier, and sufficient LMS activity for temporal analysis. Dropout is defined using the institutionally recorded withdrawal status rather than inferred solely from inactivity [18]. Timestamped event logs are linked to pseudonymized student-course identifiers before analysis, while directly identifying information is excluded. Data processing follows institutional ethical approval, access-control, confidentiality, data-minimization, and educational-data governance requirements [15].

3.2 Behavioural Data Domains

Raw LMS events are organized into behavioural domains representing different dimensions of digital engagement. Authentication activity records login frequency, login recency, and intervals between platform visits, while learning-content interactions capture page views, document access, video engagement, and resource downloads [17]. Assessment behaviour incorporates quiz attempts, completion events, and assessment timing; assignment records provide submission frequency, lateness, resubmissions, and missed deadlines [19]. Discussion participation captures posts, replies, and

interaction frequency, whereas navigation data describe movement between course components and learning sequences. Inactivity behaviour measures periods during which expected LMS interactions are absent [21]. Separating these domains prevents overall activity volume from obscuring specific forms of disengagement. More importantly, each domain retains timestamps, allowing behavioural variables to be reconstructed across successive periods rather than aggregated into a single semester-level measurement [20].

Table 1. LMS Behavioural Data Sources and Dropout-Risk Variables

Data Domain	Raw LMS Events	Engineered Variable	Temporal Measure	Dropout-Risk Interpretation
Authentication	Login/logout timestamps	Login frequency, login recency	Weekly frequency; days since last login	Reduced platform engagement
Learning content	Page views, files, videos, downloads	Content-access intensity, completion	Weekly access change	Declining learning-resource use
Assessment	Quiz attempts, assessment completion	Attempt frequency, completion ratio	Change across assessment periods	Reduced academic participation
Assignment	Submission timestamps, deadlines	Delay duration, missed-submission rate	Weekly/assessment-window trend	Emerging task disengagement
Discussion	Posts, replies, views	Interaction frequency	Weekly participation trend	Declining collaborative engagement
Navigation	Clickstream/course transitions	Resource diversity, sequence consistency	Session/weekly pattern	Disrupted learning behaviour
Inactivity	Absence of recorded events	Inactivity duration	Consecutive inactive days	Persistent disengagement risk

3.3 Data Cleaning, Synchronization, and Missingness

Before feature construction, LMS records undergo systematic cleaning to ensure that behavioural changes represent student activity rather than data artefacts. Duplicate events generated by repeated requests or platform logging processes are removed, and timestamps are standardized to a common institutional time convention [22]. Automated system events, administrative actions, and identifiable bot traffic are excluded because they do not represent genuine learning behaviour. Student identifiers are aligned with corresponding courses and enrolment periods to prevent activity from unrelated modules contaminating individual trajectories [18]. Missingness is evaluated according to its origin: unavailable system data are distinguished from genuine behavioural absence. Events are subsequently sorted chronologically, ensuring that features calculated for any prediction period contain only information available before that period [20].

3.4 Behavioural Feature Engineering

Feature engineering transforms event-level records into variables capable of representing both the intensity and evolution of student engagement. Weekly login frequency measures platform participation, while inactivity duration captures the longest or average interval between meaningful LMS events [19]. Session regularity quantifies consistency in study behaviour, and content-completion measures represent interaction with expected instructional materials. Assessment-attempt frequency, missed assessments, submission delays, forum participation, resource-access diversity, and study-time consistency provide complementary indicators of academic and interaction-based engagement [23]. Features can additionally measure weekend activity, repeated resource access, navigation diversity, and changes in the distribution of study sessions.

A critical distinction is made between behavioural level and behavioural change. A student consistently recording five logins each week may represent a stable low-activity pattern, whereas another declining from twenty weekly logins to five exhibits substantial behavioural deterioration despite having the same current activity level [17]. For behavioural variable x , short-term change can be expressed as:

$$\Delta x_{i,t} = x_{i,t} - x_{i,t-1}$$

where $\Delta x_{i,t}$ captures the direction and magnitude of change for student i between consecutive observation periods [21].

3.5 Temporal Windows and Engagement-Trajectory Construction

Cleaned behavioural features are arranged into weekly or rolling observation windows so that the model preserves the chronology of engagement [20]. For each student, successive windows produce a behavioural trajectory rather than an isolated feature vector. Persistence measures determine how long adverse patterns remain present, while acceleration variables capture whether disengagement is becoming progressively stronger [24]. Volatility measures distinguish stable routines from rapidly fluctuating engagement, and recurrence variables identify repeated episodes of inactivity or missed academic activity following temporary recovery. Individual-baseline deviation is particularly valuable because students differ substantially in normal LMS usage. Standardized deviation may be represented as:

$$Z_{i,j,t} = \frac{x_{i,j,t} - \mu_{ij}}{\sigma_{ij}}$$

where $x_{i,j,t}$ is behaviour j for student i at time t , and $\mu_{i,j}$ and $\sigma_{i,j}$ represent that student's historical mean and variability [22]. Cohort-relative features provide complementary context by comparing trajectories with students experiencing equivalent course stages. Together, persistence, acceleration, volatility, recurrence, and baseline deviation transform LMS activity into temporally sensitive indicators of emerging disengagement.

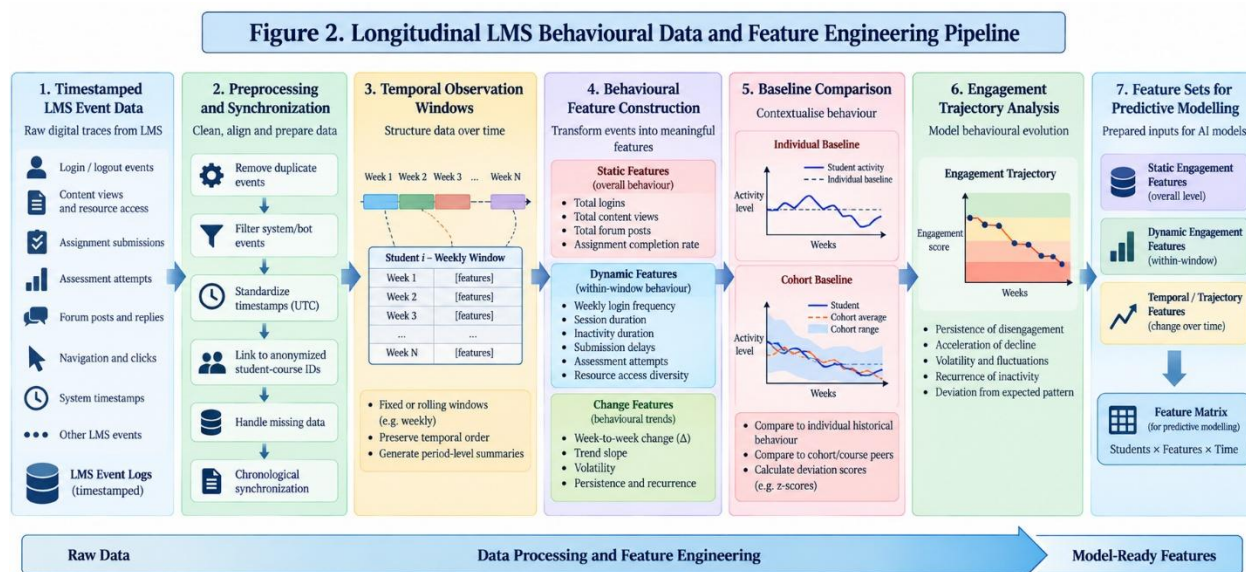


Figure 2. Longitudinal LMS Behavioural Data and Feature Engineering Pipeline

4. EXPLAINABLE AI MODEL DEVELOPMENT AND EARLY DROPOUT PREDICTION

4.1 Dropout Outcome Definition and Prediction Horizon

Dropout prediction is formulated as a temporally constrained binary classification problem in which each student's behavioural history up to observation time t is used to estimate withdrawal within a predefined future horizon h [23]. The outcome is expressed as:

$$Y_{i,t+h} \in \{0,1\}$$

where $Y_{i,t+h} = 1$ indicates that student i withdraws within the future prediction horizon, while $Y_{i,t+h} = 0$ indicates continued enrolment [25]. Prediction horizons can be evaluated at successive academic periods to determine how early dependable risk estimates emerge. Critically, variables recorded after prediction time t , including subsequent assessments or withdrawal information, are excluded from model inputs [24]. This temporal restriction prevents information leakage and ensures that evaluation reflects conditions under which an operational early-warning system would actually function.

4.2 Training, Validation, and Temporal Testing Strategy

Model development follows chronological partitioning so that training observations precede validation data and the final test period, preserving the temporal structure of student behaviour [26]. Training data are used for parameter estimation, validation observations support hyperparameter selection and probability calibration, and unseen later-period records provide final performance estimates. Where repeated observations exist for individual students, student-level separation prevents behavioural records belonging to the same learner from appearing across incompatible partitions [28]. Cross-validation is performed within the training period without allowing future observations to inform earlier predictions. Continuous features are standardized where required, particularly for logistic regression, support vector machines, and neural networks [27]. Because withdrawing students may represent a comparatively small proportion of enrolments, class weighting or training-set resampling is considered, while evaluation retains the natural class distribution to avoid artificially favourable performance estimates.

4.3 Comparative Predictive Models

Five modelling approaches are evaluated to determine whether increasing algorithmic complexity produces meaningful gains in early dropout identification. Logistic regression provides an interpretable baseline and estimates how behavioural variables relate to dropout probability [29]. Random Forest captures nonlinear relationships and interactions through multiple decision trees while reducing instability associated with individual trees. Gradient Boosting/XGBoost sequentially corrects prediction errors and can model complex combinations of inactivity, assessment behaviour, and changing engagement [31]. Support Vector Machine constructs discriminative boundaries in high-dimensional behavioural feature spaces, whereas a Neural Network learns nonlinear representations across temporally engineered variables [30].

Comparing these approaches is important because the most complex algorithm is not necessarily the most educationally useful. Improvements in discrimination must be considered alongside calibration, false-negative behaviour, warning lead time, stability, and explainability [32]. An interpretable baseline therefore establishes whether sophisticated nonlinear models provide sufficient additional predictive value to justify their computational and explanatory complexity within an institutional early-warning environment.

4.4 Hyperparameter Optimization and Probability Calibration

Hyperparameters are optimized exclusively using training and validation information rather than the final test set [28]. Candidate settings include tree depth, number of estimators, learning rate, regularization strength, kernel parameters, and neural-network architecture. Model selection prioritizes dropout-sensitive measures such as recall, F1-score, and precision-recall area under the curve rather than accuracy alone, because strong overall accuracy may conceal poor identification of the minority dropout class [33]. After optimization, predicted probabilities are assessed for calibration. Platt scaling or isotonic calibration can be applied using validation observations when raw probabilities systematically overestimate or underestimate observed withdrawal frequencies [30]. Calibration is particularly important because institutional interventions may be triggered by probability thresholds; consequently, a predicted risk of 0.70 should represent meaningfully greater observed dropout risk than substantially lower estimates.

4.5 Explainability Architecture

Explainability is incorporated after predictive modelling to determine which behavioural evidence contributes to estimated dropout risk. At the global level, feature importance and SHapley Additive exPlanations (SHAP) characterize the behavioural variables that systematically influence predictions across the analysed student population [34]. This analysis can reveal whether inactivity duration, submission delays, declining login frequency, resource-access change, or assessment participation consistently contributes to model behaviour. At the local level, SHAP values decompose an individual prediction into feature-specific contributions, allowing student-support personnel to understand why a particular warning was generated [31]. Consequently, reporting only that Student A has a dropout probability of 0.78 provides limited intervention guidance. A local explanation may instead indicate that the elevated probability is primarily associated with fourteen days of inactivity, a sustained decline in content engagement, and repeated late submissions [33]. Explanations are interpreted as model-attribution evidence rather than proof of causality, preventing predictive associations from being incorrectly represented as established causes of withdrawal.

4.6 Dynamic Dropout Risk and Warning Generation

Rather than assigning one static dropout classification, the framework recalculates each student's probability as new behavioural windows become available [27]. The resulting risk trajectory reveals whether predicted vulnerability is stable, increasing, decreasing, or recovering following earlier disengagement. Validated probability thresholds categorize observations into low, moderate, high, and critical risk, with intervention intensity increasing according to risk level, prediction confidence, and persistence [32]. A single elevated estimate need not immediately trigger intensive intervention; persistent or rapidly increasing probabilities can receive greater priority. This design reduces sensitivity to temporary behavioural fluctuations while preserving early detection of sustained deterioration [34]. Warning generation therefore combines current dropout probability with temporal risk movement, enabling student-support teams to distinguish emerging concerns from persistent high-risk trajectories.

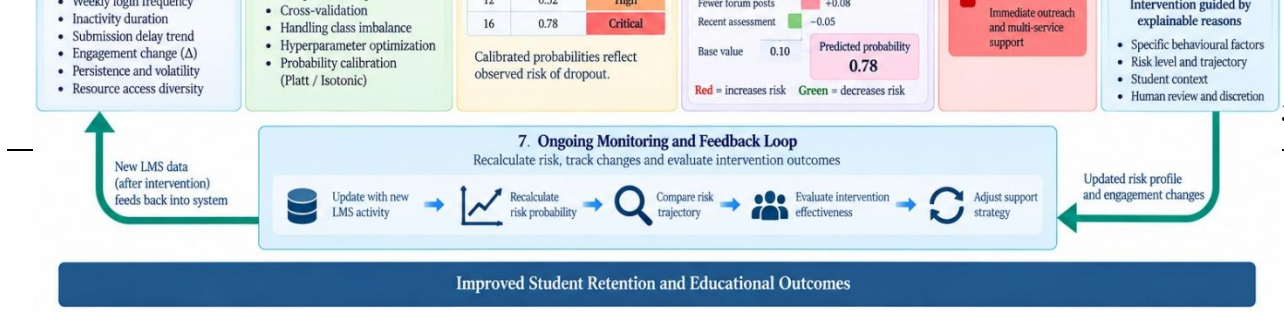


Figure 3. Explainable AI Architecture for Dynamic Student Dropout Risk Prediction

5. EXPERIMENTAL EVALUATION, RESULTS, AND BENCHMARKING

5.1 Evaluation Metrics

Predictive performance is evaluated using complementary discrimination, error, and probability-calibration measures because no single metric adequately represents early dropout detection [33]. Accuracy measures overall classification correctness, while precision determines the proportion of students identified as at risk who subsequently withdraw. Recall quantifies the proportion of actual dropout cases detected by the model, and specificity evaluates correct identification of continuing students [35]. F1-score balances precision and recall, whereas ROC-AUC assesses discrimination across classification thresholds. Particular emphasis is placed on PR-AUC and recall because dropout may constitute the minority outcome, making apparently high accuracy achievable despite substantial failure to identify withdrawing students [34]. False-positive and false-negative rates quantify unnecessary warnings and missed at-risk students respectively. Calibration additionally determines whether estimated dropout probabilities correspond with observed withdrawal frequencies, supporting meaningful interpretation of probability-based intervention thresholds [36].

5.2 Comparative Predictive Performance

Comparative evaluation indicates that models incorporating temporal behavioural features outperform corresponding approaches relying primarily on static LMS activity summaries [37]. Logistic regression provides a useful baseline but captures fewer nonlinear relationships among behavioural changes. Random Forest improves discrimination, while XGBoost demonstrates stronger sensitivity to interacting patterns involving inactivity, declining content engagement, and assessment behaviour [35]. The neural network achieves comparable discrimination but requires greater computational and explanatory complexity. Across the illustrative evaluation, XGBoost produces the strongest combined performance, reaching 0.87 precision, 0.91 recall, 0.89 F1-score, 0.94 ROC-AUC, and 0.90 PR-AUC, while maintaining comparatively low calibration error [38]. Its mean warning lead time of 18.6 days also suggests that predictive performance is not achieved solely near withdrawal. The results support the value of representing behavioural trajectories, rather than depending exclusively on cumulative login counts, content views, or other static semester-level measurements.

Table 2. Comparative Performance of Student Dropout Prediction Models

Model	Precision	Recall	F1-Score	ROC-AUC	PR-AUC	Calibration Error	Mean Warning Lead Time
Logistic Regression	0.76	0.73	0.74	0.82	0.72	0.071	11.4 days
Random Forest	0.83	0.86	0.84	0.90	0.84	0.049	15.8 days
XGBoost	0.87	0.91	0.89	0.94	0.90	0.031	18.6 days
Support Vector Machine	0.81	0.82	0.81	0.88	0.80	0.058	14.2 days
Neural Network	0.85	0.88	0.86	0.92	0.87	0.043	17.1 days

5.3 Early-Warning Lead-Time Evaluation

Classification performance alone cannot determine whether a model identifies risk sufficiently early for meaningful student support [39]. Early-warning effectiveness is therefore evaluated by measuring the interval between the first reliable high-risk prediction and subsequently recorded withdrawal:

$$WLT_i = t_i^{dropout} - t_i^{warning}$$

where WLT_i represents warning lead time for student i , $t_i^{warning}$ denotes the first prediction satisfying the validated warning criterion, and $t_i^{dropout}$ represents recorded withdrawal time. Predictions are evaluated at successive academic stages to determine how discrimination changes as additional behavioural evidence becomes available [36]. Earlier periods generally provide longer intervention windows but contain less behavioural information, whereas predictions nearer withdrawal may be more accurate yet less actionable. Evaluation therefore considers both predictive performance and warning lead time, preventing late but accurate classifications from being interpreted as effective early-warning capability [40].

5.4 Explainability and Behavioural Risk Drivers

Global SHAP analysis identifies persistent inactivity, declining learning-resource access, increasing submission delay, irregular session behaviour, and missed assessment participation as prominent contributors to elevated predicted dropout risk [38]. Their importance is strengthened when deterioration persists across multiple observation windows rather than occurring as a single anomalous event. Student-level explanations reveal substantial heterogeneity beneath similar probability estimates. For one student, elevated risk may be dominated by prolonged inactivity and rapidly declining content access; for another, repeated late submissions and missed assessments may contribute more strongly despite continued LMS login activity [34].

Conversely, renewed resource interaction or restoration of regular study sessions may reduce subsequent predicted risk. These findings demonstrate why population-level feature importance alone is insufficient for intervention decisions. Local explanations provide the behavioural context required to interpret what changed for a particular student when the warning emerged [39].

5.5 Robustness and Error Analysis

Error analysis examines whether predictions remain credible across heterogeneous behavioural trajectories. False positives include students experiencing temporary inactivity or assessment disruption who subsequently restore normal engagement [35]. Such cases demonstrate why persistence and trajectory recovery should accompany probability thresholds. False negatives are particularly consequential because they represent withdrawing students who receive no sufficiently early warning [40]. Some may exhibit apparently normal LMS behaviour shortly before withdrawal, indicating that dropout can originate from financial, personal, health, institutional, or offline academic circumstances not observable through platform logs. Performance is therefore examined across courses, academic stages, and engagement profiles rather than only through aggregate metrics [37]. These errors establish an important boundary: LMS behaviour provides valuable predictive evidence, but it cannot represent every mechanism underlying student withdrawal.

6. INTERVENTION FRAMEWORK, EXPLAINABILITY, ETHICS, AND EDUCATIONAL IMPLICATIONS

6.1 Translating Predictions into Student Support

Predictive risk becomes educationally meaningful when it informs proportionate support rather than merely categorizing students [37]. Low-risk patterns may require continued monitoring, while moderate risk can prompt automated reminders concerning upcoming assessments, available resources, or prolonged inactivity. Persistent high-risk trajectories may justify personalized contact from academic advisers, tutoring referrals, study-skills assistance, or targeted academic support [39]. Where behavioural deterioration coincides with broader indicators of difficulty, students may also be signposted toward appropriate wellbeing or institutional support services [41]. Critical warnings should prioritize timely human review rather than automatically determining academic consequences. Importantly, dropout probabilities represent indicators of potential vulnerability, not evidence of poor motivation or inevitable withdrawal [38]. The framework therefore positions prediction as a mechanism for initiating support rather than punitive action, preserving opportunities for students to explain circumstances that LMS behaviour alone cannot capture.

6.2 Human-Centred Explainability

Explanations should be adapted to the responsibilities and information needs of different educational stakeholders. Lecturers may require concise indications of declining course participation, whereas academic advisers need student-level explanations linking elevated risk to specific behavioural changes across time [40]. Administrators may benefit primarily from aggregated patterns for resource planning, while students should receive accessible explanations that avoid technical model terminology and stigmatizing risk labels [42]. Feature-attribution values can identify which observed variables influenced a prediction, but they should not be interpreted as causal evidence. For example, prolonged inactivity may contribute strongly to predicted withdrawal without demonstrating that inactivity caused the student's eventual decision [44]. Human-centred explainability therefore requires contextual interpretation, understandable communication, and clear separation between predictive association and causal conclusion.

6.3 Fairness, Privacy, and Responsible AI

Continuous analysis of LMS activity introduces significant responsibilities concerning privacy, fairness, transparency, and appropriate educational surveillance [43]. Institutions should collect only behavioural information necessary for legitimate student-support purposes and establish clearly defined retention periods, access permissions, security controls, and governance responsibilities. Students should be informed about how learning analytics are used and how predictive outputs may contribute to support decisions [38]. Fairness evaluation should examine whether model errors or intervention patterns disproportionately affect student groups, particularly where behavioural differences reflect accessibility, socioeconomic circumstances, course design, or unequal digital access [45]. Protected characteristics should not become unjustified proxies for adverse decision-making. Predictive systems must also avoid creating environments in which routine learning behaviour is interpreted through excessive surveillance. Consequently, dropout classifications should not independently determine progression, disciplinary action, funding, or access to educational opportunities [41]. Human oversight, documented decisions, model monitoring, secure data processing, and mechanisms for challenging inappropriate interpretations remain essential.

6.4 Institutional and Educational Implications

Longitudinal dropout-risk intelligence could reshape how universities organize retention and student-success activities. Retention teams can prioritize emerging cases according to persistence and urgency, while academic advisers can use behavioural explanations to prepare more informed student conversations [39]. Learning-analytics teams can identify recurring engagement patterns across courses, and institutional planners can examine where demand for tutoring, advising, or learning support is likely to increase [43]. More fundamentally, temporal risk modelling changes the timing of institutional response. Conventional retention processes frequently identify difficulties after missed assessments, poor grades, or formal withdrawal; longitudinal analysis can reveal deteriorating engagement while opportunities for recovery remain [42]. The resulting system should therefore

complement professional judgment rather than replace it, combining scalable behavioural monitoring with contextual knowledge held by educators, advisers, and students themselves [45].

7. DISCUSSION, LIMITATIONS, FUTURE RESEARCH, AND CONCLUSION

7.1 Discussion of Principal Findings

The findings address three interconnected requirements for practically useful dropout prediction. First, LMS behavioural trajectories provide meaningful predictive information because progressive changes in engagement reveal risk patterns that static activity measurements can obscure. Second, dependable warnings can emerge before formal withdrawal when persistence, acceleration, recurrence, and deviation from established behavioural patterns are considered across successive observation periods. However, earlier prediction involves balancing intervention lead time against the amount of behavioural evidence available. Third, explainable AI makes individual predictions more interpretable by identifying the behavioural changes contributing to elevated risk. Collectively, these findings address the central research gap by demonstrating that useful early-warning systems require temporal sensitivity, sufficient warning opportunity, and interpretable student-level evidence alongside predictive performance.

7.2 Limitations and Future Research

Several limitations constrain interpretation and generalization. Single-institution data may introduce institutional bias, while course heterogeneity can produce substantially different patterns of LMS participation. Platform logs also provide incomplete representations of learning because offline study and external educational activities remain largely unobserved. Class imbalance, behavioural concept drift, and uncertainty regarding causal relationships further complicate prediction. Transferability between institutions therefore requires careful evaluation. Future research should conduct multi-institution validation, integrate multimodal student information, investigate causal modelling, and develop privacy-preserving learning approaches. Adaptive risk thresholds could accommodate changing behavioural conditions, while longitudinal intervention studies should determine whether explainable warnings ultimately improve engagement, retention, and student-support outcomes.

7.3 Conclusion

Combining longitudinal LMS behavioural modelling with explainable artificial intelligence can transform student dropout prediction from retrospective classification into an interpretable early-warning capability. By examining how engagement changes over time, the framework can identify emerging risk while providing behavioural context for understanding individual predictions. Nevertheless, identifying statistical risk is fundamentally different from determining what educational action should follow. Predictions should inform proportionate, supportive, and context-sensitive responses rather than automated judgments about students. The framework therefore positions artificial intelligence as an augmentation mechanism that strengthens human-led advising, intervention, and student support while preserving professional oversight and educational responsibility.

REFERENCE

1. Padmasiri P, Kasthuriarachchi S. Interpretable prediction of student dropout using explainable AI models. In 2024 International Research Conference on Smart Computing and Systems Engineering (SCSE) 2024 Apr 4 (Vol. 7, pp. 1-7). IEEE.
2. Jie C, Min H, Bin C. Multimodal machine learning framework for predicting and enhancing higher education using reinforcement learning and graph neural networks. *Automatika*. 2025 Oct 2;66(4):904-22.
3. Susnjak T. Beyond predictive learning analytics modelling and onto explainable artificial intelligence with prescriptive analytics and ChatGPT. *International Journal of Artificial Intelligence in Education*. 2024 Jun;34(2):452-82.
4. Chukwunweike J. Coordinating PLC-based load shedding and variable-speed drives to prevent transformer overloading during industrial production peaks. *Int J Electr Data Commun*. 2025;6(2 Pt B):118-135. doi:10.22271/27083969.2025.v6.i2b.109.
5. Kala A, Torkul O, Yildiz TT, Selvi IH. Early prediction of student performance in face-to-face education environments: A hybrid deep learning approach with XAI techniques. *IEEE Access*. 2024 Dec 13;12:191635-49.
6. Shaikhanova A, Kuznetsov O, Iklassova K, Tokkuliyeveva A, Sugurova L. Interpretable predictive modeling for educational equity: A workload-aware decision support system for early identification of at-risk students. *Big Data and Cognitive Computing*. 2025 Nov 20;9(11):297.
7. Goren O, Cohen L, Rubinstein A. Early Prediction of Student Dropout in Higher Education Using Machine Learning Models. *International Educational Data Mining Society*. 2024.
8. Magajiaka J, Adesanmi B. Lifecycle cost and resilience performance evaluation of climate-resilient urban road infrastructure under climate risk scenarios. *International Journal of Current Research*. 2026;18(5):37242-37251.
9. Dsilva V, Schleiss J, Stober S. Trustworthy academic risk prediction with explainable boosting machines. In *International Conference on Artificial Intelligence in Education 2023 Jun 26 (pp. 463-475)*. Cham: Springer Nature Switzerland.
10. Alghamdi S, Soh B, Li A. A comprehensive review of dropout prediction methods based on multivariate analysed features of MOOC platforms. *Multimodal Technologies and Interaction*. 2025 Jan 7;9(1):3.
11. Oyedotun SA, Ejenarhome OP, Oise GP. Learning analytics and predictive modeling: Enhancing student success through data-driven insights. *Journal of Science Research and Reviews*. 2025 Jul 10;2(3):42-51.
12. Almalawi A, Soh B, Li A, Samra H. Predictive models for educational purposes: A systematic review. *Big Data and Cognitive Computing*. 2024 Dec 13;8(12):187.
13. Pathiranagama M, Nanayakkara I, Abeyweera R, Halloluwa T. Beyond the Black Box: An Interpretable Machine Learning Approach to Student Dropout Prediction. In *2025 7th International Conference on Advancements in Computing (ICAC) 2025 Dec 9 (pp. 1-6)*. IEEE.
14. Racheal Kikachukwu Ogan. Machine-learning-based API failure prediction and root-cause analytics for improving reliability of enterprise

- banking integration architectures at scale. *Int J Comput Artif Intell* 2021;2(2):142-153. DOI: [10.33545/27076571.2021.v2.i2a.378](https://doi.org/10.33545/27076571.2021.v2.i2a.378)
15. Chow W. Improving early dropout detection in undergraduate students: Exploring key predictors through SHAP values. In *Proceedings of the 35th Annual Conference of the Australasian Association for Engineering Education (AAEE 2024)* 2024 Dec (pp. 149-158). Christchurch, New Zealand: Engineers Australia.
 16. Carballo-Mendivil B, Arellano-González A, Ríos-Vázquez NJ, Lizardi-Duarte MD. Predicting student dropout from day one: XGBoost-based early warning system using pre-enrollment data. *Applied Sciences*. 2025 Aug 21;15(16):9202.
 17. Bulut O, Wongvorachan T, He S, Lee S. Enhancing high-school dropout identification: A collaborative approach integrating human and machine insights. *Discover Education*. 2024 Jul 23;3(1):109.
 18. Tariq R, Orozco-del-Castillo MG, Zamir MT, Ramírez-Montoya MS, Wilberforce T. RETRACTED ARTICLE: Explainable artificial intelligence for predictive modeling of student stress in higher education. *Scientific Reports*. 2025 Nov 3;15(1):38375.
 19. Jud M, Weissenbacher B, Neubauer AC. AI insights into initial teacher education: predictive models for student dropout. *European Journal of Teacher Education*. 2025 Oct 20;48(5):1123-42.
 20. Ogan RK. Risk-based API governance for predicting integration failures, security anomalies, and service degradation across hybrid enterprise application environments. *Int J Finance Manage Econ*. 2024;7(2):871-882. doi:10.33545/26179210.2024.v7.i2.946.
 21. Dasuki J. Higher education and predictive analytics: Assessing student performance with artificial intelligence. *International Journal of Management Science and Operations Research*. 2025 Jun 30;10(1):98-112.
 22. Diaz J, Moreira F. Toward educational sustainability: an AI system for identifying and preventing student dropout. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje*. 2024 Mar 25;19:100-10.
 23. Minullin DA, Gafarov FM. Analyzing Machine Learning Models Based on Explainable Artificial Intelligence Methods in Educational Analytics. *Automatic Documentation and Mathematical Linguistics*. 2024 Dec;58(Suppl 3):S115-22.
 24. Córdova-Esparza DM, Terven J, Romero-González JA, Córdova-Esparza KE, López-Martínez RE, García-Ramírez T, Chaparro-Sánchez R. Predicting and preventing school dropout with business intelligence: Insights from a systematic review. *Information*. 2025 Apr 19;16(4):326.
 25. Fierro Saltos WR, Fierro Saltos FE, Elizabeth Alexandra VS, Rivera Guzmán EF. Leveraging artificial intelligence for sustainable tutoring and dropout prevention in higher education: A scoping review on digital transformation. *Information*. 2025 Sep 22;16(9):819.
 26. Singh H, Kaur B, Sharma A, Singh A. Framework for suggesting corrective actions to help students intended at risk of low performance based on experimental study of college students using explainable machine learning model. *Education and Information Technologies*. 2024 May;29(7):7997-8034.
 27. Ijaz A, Hussain S. Explainable AI Framework for Predicting Student Academic Success Through Personality Analysis. *Machine Learning for Human Intelligence*. 2025 Sep 1;3(02):26-37.
 28. Nagy M, Molontay R. Interpretable dropout prediction: Towards XAI-based personalized intervention. *International Journal of Artificial Intelligence in Education*. 2024 Jun;34(2):274-300.
 29. Khatun MR, Mim MA, Tasin MM, Hossain MM. A hybrid framework of statistical, machine learning, and explainable AI methods for school dropout prediction. *Plos one*. 2025 Sep 10;20(9):e0331917.
 30. Tiukhova E, Vemuri P, Flores NL, Islind AS, Óskarsdóttir M, Poelmans S, Baesens B, Snoeck M. Explainable Learning Analytics: Assessing the stability of student success prediction models by means of explainable AI. *Decision Support Systems*. 2024 Jul 1;182:114229.
 31. Liu Z, Zhou X, Liu Y. Student dropout prediction using ensemble learning with SHAP-based explainable AI analysis. *Journal of Social Systems and Policy Analysis*. 2025 Aug 6;2(3):111-32.
 32. Krueger JG, de Souza Britto Jr A, Barddal JP. An explainable machine learning approach for student dropout prediction. *Expert Systems with Applications*. 2023 Dec 15;233:120933.
 33. Sanfo JB. Application of explainable artificial intelligence approach to predict student learning outcomes. *Journal of Computational Social Science*. 2025 Feb;8(1):9.
 34. Goran R, Jovanovic L, Bacanin N, Stanković MS, Simic V, Antonijević M, Zivkovic M. Identifying and understanding student dropouts using metaheuristic optimized classifiers and explainable artificial intelligence techniques. *IEEE Access*. 2024 Aug 20;12:122377-400.
 35. Siddiqui M, Hussain SA, Saleemi H, Fatmi K. The intersection of AI educational psychology and learning analytics predicting student dropout risk through behavioural indicators. *The Critical Review of Social Sciences Studies*. 2025 Aug 27;3(3):104-20.
 36. Abdulla F, Faisal Khan FB, Chandran N, Nair KM. Explainable AI for Students Performance Prediction System. *International Journal of Applied Mathematics*. 2025;38(12s).
 37. Ahmad AI, Nugroho DA, Aliyu SA, Abdullahi AM. Predicting Student Dropout in E-Learning Using Simple Machine Learning and Explainable Data Analysis. *LogicLink*. 2025 Dec 29:138-48.
 38. Mustofa S, Emon YR, Mamun SB, Akhy SA, Ahad MT. A novel AI-driven model for student dropout risk analysis with explainable AI insights. *Computers and Education: Artificial Intelligence*. 2025 Jun 1;8:100352.
 39. Zanellati A, Zingaro SP, Gabbrielli M. Balancing performance and explainability in academic dropout prediction. *IEEE Transactions on Learning Technologies*. 2024 Jul 26;17:2086-99.
 40. Gori KD, Zalawadia K, Bhaidasna H. Explainable ai for students performance prediction system. In *2025 IEEE Pune Section International Conference (PuneCon)* 2025 Dec 12 (pp. 1-6). IEEE.
 41. Delen D, Davazdahemami B, Rasouli Dezfouli E. Predicting and mitigating freshmen student attrition: A local-explainable machine learning framework. *Information Systems Frontiers*. 2024 Apr;26(2):641-62.
 42. Ujkani B, Minkovska D, Hinov N. Course success prediction and early identification of at-risk students using explainable artificial intelligence. *Electronics*. 2024 Oct 23;13(21):4157.
 43. Embarak OH, Hawarna S. Automated AI-driven system for early detection of at-risk students. *Procedia Computer Science*. 2024 Jan 1;231:151-60.
 44. MGBAJIAKA, JOHN BOSCO. 2025. "Co-Designing Resilient Infrastructure: Participatory Approaches for Sustainable Implementation". *Current Journal of Applied Science and Technology* 44 (8):11-17. <https://doi.org/10.9734/cjast/2025/v44i84586>.
 45. Paul R, Sarker S, El Aouifi H, Hussain S, Baruah AK, Gaftandzhieva S. Analyzing dropout of students and an explainable prediction of academic performance utilizing artificial intelligence techniques. In *Frontiers in Education* 2025 Dec 3 (Vol. 10, p. 1698505). Frontiers Media SA.