

Article

A Closed-Loop Artificial Intelligence System for Process-Oriented Student Assessment and Early Performance Prediction in Higher Education

Truong Thi Huong Giang ¹ and Young-Jae Ryoo ^{2,*}

¹ Department of Information Technology, Tay Nguyen University, Daklak 632090, Vietnam; tthgiang@ttn.edu.vn

² Department of Electrical Engineering, Mokpo National University, Muan 58554, Jeonnam, Republic of Korea

* Correspondence: yjryoo@mnu.ac.kr

Abstract

The increasing amount of detailed learning data in higher education has shifted assessment methods from traditional scores to continuous assessment of student learning performance. This paper presents an applied closed-loop AI system that supports early performance prediction, informed risk decision-making, and personalized intervention planning. The system utilizes aggregated weekly data from learning management systems, including self-regulated learning metrics and participation indicators. A supervised regression model predicts near-future learning performance while employing chronological training strategies to prevent information leakage. Emphasizing decision-driven assessment, the system translates predictive outputs into a resource-sensitive decision framework to identify students requiring early support. Explainable artificial intelligence (XAI) techniques enhance the model's transparency, providing insights into key learning factors and individual student-level explanations. The system generates actionable recommendations based on these insights. Its closed-loop architecture integrates data collection, prediction, interpretation, recommendation generation, and feedback logging, allowing instructors to monitor learning pathways and examine subsequent learning outcomes. Rather than claiming causal intervention effects, the proposed system illustrates how established AI, explainable learning analytics, and decision-support techniques can be operationally integrated to support ongoing student assessment in higher education.

Keywords: artificial intelligence in higher education; learning analytics; early performance prediction; explainable AI; decision-oriented analytics; closed-loop systems

Academic Editors: María Lucia Barrón-Estrada and Ramón Zatarain-Cabada

Received: 2 June 2026

Revised: 1 July 2026

Accepted: 8 July 2026

Published: 9 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Artificial intelligence (AI) is increasingly transforming higher education through data-driven practices, which include personalized learning experiences and near-real-time support for pedagogical decision-making. As online and blended learning environments become more widespread, the need for early warning systems and timely, evidence-based feedback has led to a crucial area of research and practice known as Learning Analytics (LA) and process-oriented assessment [1–4]. Despite significant advancements, there remains a notable gap between having an “accurate predictive model” and imple-

menting “an operational system that can effectively support pedagogical actions.” Many existing solutions only provide risk scores without fully integrating model explainability, concrete intervention recommendations, impact measurement, or continuous updating cycles [1,2,5–8].

Recent advances in artificial intelligence have further enhanced the capability of learning analytics systems by integrating predictive modeling, adaptive feedback mechanisms, and intelligent decision support. In particular, recent studies have explored data-driven approaches for early warning systems, personalized learning pathways, and real-time intervention strategies in higher education environments [9–11]. These systems leverage large-scale behavioral data from Learning Management Systems (LMS) to provide actionable insights for both instructors and learners, moving beyond descriptive analytics toward prescriptive and decision-oriented analytics.

In the context of learning analytics, a closed-loop system refers to a cycle where data collection, analysis, intervention, and evaluation are closely integrated and consistently connected. If only the analysis and assessment of the learning process are conducted without making informed pedagogical decisions and timely interventions for learners, then the analysis becomes ineffective [8]. Closed-loop learning analytics emphasizes the ongoing relationship between analytical outcomes and pedagogical actions, followed by systematic monitoring of their effects. Recent research outlines this process as a series of stages, including data collection, modeling and interpretation, targeted or supportive interventions, and refinement based on feedback from both the model and practical application. This approach allows learning analytics to serve as an active component of educational decision-making rather than merely a passive reporting tool [6,12]. Figure 1 illustrates the conceptual closed-loop learning analytics process as described in [6], in which analytical outcomes are linked to pedagogical interventions and continuously refined through feedback.

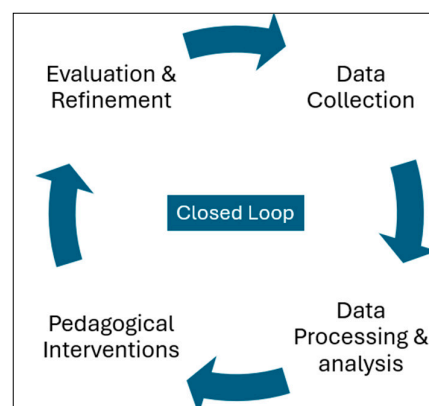


Figure 1. The closed-loop learning analytics framework adapted from [6].

This study builds on the concept of closed-loop learning analytics and aims to operationalize it as a practical system in the context of higher education. While previous research has highlighted the importance of connecting analytics to teaching actions and feedback, fewer studies have shown how this connection can be implemented in a realistic, decision-oriented, and understandable way within everyday academic functions.

In this work, we develop the closed-loop framework into an AI-enabled system that operates on a week-by-week basis. This system links learning data from an institutional Learning Management System (LMS) to short-term performance predictions, helps inform pedagogical decision-making, and assists in planning actionable interventions. Instead of treating the closed-loop process as an abstract concept, our proposed system

clearly defines the necessary data granularity, modeling strategies, decision policies, and feedback mechanisms. These elements are crucial for supporting continuous instructional adjustments at both course and class levels. This operational approach allows learning analytics to progress beyond theoretical alignment and support feedback-informed pedagogical decision-making in real educational environments.

To translate weekly predictions into effective teaching strategies, the recommendation system follows two key design principles that emphasize the practical use of decisions rather than just predictive accuracy. First, the predictions are decision-driven, providing continuous performance scores on a standardized scale from 0 to 100, instead of simple binary risk labels. This approach allows for more adaptable decision-making, enabling faculty and academic support units to identify students who may require support through either rank-based selection or threshold-based strategies, tailored to their available resources and teaching constraints.

Secondly, the system establishes a clear connection between interpretation and action through integrated explainable artificial intelligence (XAI) mechanisms. It generates model interpretations at both global and local levels to identify key learning indicators and diagnose specific student factors that contribute to predicted performance declines in a given week. These interpretations are then used to create feasible, short-term recommendations, outlining possible behavioral adjustments for the next teaching cycle. By incorporating interpretability into the decision-making process, the system ensures that predicted outcomes are understandable, actionable, and educationally sound, rather than simply presenting vague risk scores.

The use of AI-supported decision-making in education also raises important concerns regarding transparency, human oversight, and responsible implementation. Recent work on generative AI in universities and academic assessment highlights the need for clear institutional practices, ethical usage guidelines, monitoring mechanisms, and instructor-led decision processes [13–15]. Accordingly, the proposed system treats predicted scores, explanations, and recommendations as decision-support outputs that require instructor review rather than as automatic intervention or assessment decisions.

This issue becomes particularly important in GenAI-aware assessment contexts. Recent studies on GenAI in higher education indicate that institutional policies, assessment design, academic integrity practices, and human oversight strongly influence how AI-supported tools should be implemented in universities [13,14,16]. In courses where GenAI is permitted, restricted, or systematically integrated into learning activities, LMS-based indicators such as submission behavior, discussion quality, time-on-task, and assignment performance may reflect different learning processes than in conventional assessment settings. Therefore, the proposed system should not be interpreted as a tool for detecting AI-generated work or determining academic misconduct. Instead, it should be understood as a process-oriented decision-support framework that helps instructors monitor observable learning behaviors, interpret predicted performance risks, and plan reviewable pedagogical support within the assessment policy of a specific course. From an implementation perspective, universities can apply the proposed framework as a weekly decision-support workflow within LMS-mediated courses. This involves identifying reliable LMS data sources, aggregating learner activities into learner–course–week indicators, training prediction models using time-aware validation, presenting predicted scores and explanations to instructors, mapping risk-related factors to feasible support actions, and logging learner responses in subsequent weeks. However, when the framework is transferred to other disciplines, institutions, LMS environments, or GenAI-aware assessment designs, the feature set, model parameters, decision thresholds, and recommendation mappings should be recalibrated according to local pedagogical practices and assessment policies.

This study contributes to the learning analytics literature by supporting the practical implementation of closed-loop systems in higher education. First, it operationalizes the conceptual idea of closed-loop learning analytics as a deployable, week-by-week AI-enabled system. This system explicitly connects learning data, short-term performance predictions, pedagogical decision-making, and intervention feedback within regular academic practices. Second, the study applies a decision-oriented score-to-action approach that links continuous performance predictions to rank-based and utility-aware intervention strategies. This approach supports transparent and resource-sensitive instructional decisions and extends static risk labeling toward more flexible intervention planning. Third, the research illustrates how explainable artificial intelligence can be systematically integrated into the closed-loop process to bridge the gap between model interpretation and actionable pedagogical strategies. This integration links analytic explanations to practical, short-term behavioral recommendations. Collectively, these elements position learning analytics as a practical decision-support tool for pedagogy in real educational settings, while recognizing that stronger causal evidence of intervention effectiveness requires further controlled or quasi-experimental evaluation.

The remainder of this paper is organized as follows. Section 2 introduces the institutional data sources and the construction of weekly process-oriented learning indicators derived from LMS activity logs. Section 3 presents the proposed closed-loop learning analytics framework, encompassing the formulation of the prediction target, the predictive modeling approach, the decision-oriented intervention strategy, the integration of explainable artificial intelligence (XAI), and the operational feedback cycle. Section 4 reports the experimental results and discusses the effectiveness and practical implications of the proposed system in a real instructional context. Finally, Section 5 concludes the paper and outlines directions for future research.

2. Data Sources and Weekly Process-Oriented Indicators

2.1. Data Sources and Learning Activity Logs

This study utilizes data on learning behaviors gathered from a Learning Management System (LMS) in a higher education setting. The LMS supports various teaching activities, including delivering course content, submitting assignments, participating in online discussions, and conducting formative assessments. All learner interactions with the system are automatically recorded in activity logs, creating a longitudinal record of the learning process throughout the academic term.

The raw data consists of timestamped event records related to individual learners and various course components. This includes login events, content access, assignment submissions, discussion participation, and feedback interactions. These logs are generated as part of standard instructional practices and do not require any additional data collection instruments. This approach ensures ecological validity and minimizes disruptions to teaching and learning activities [17].

This study focuses on using process-oriented behavioral data to facilitate early warning systems and timely pedagogical interventions, rather than relying on static demographic information or end-of-course outcomes. This approach aligns with previous research in learning analytics, which has demonstrated that process-oriented indicators offer more actionable insights into learning trajectories. They are particularly effective for making formative decisions and providing instructional support [12].

2.2. Study Context and Dataset Characteristics

The empirical data used in this study were collected from the institutional Learning Management System during the 2025 academic year. The study context included two un-

dergraduate programming courses within the information technology curriculum: Object-Oriented Programming and Mobile Programming. The Object-Oriented Programming course included 85 participating students, while the Mobile Programming course included 60 participating students, resulting in 145 course-level student enrollments across the two courses. The courses were taught by two instructors. The observation period covered nine instructional weeks, from ISO week 40 to ISO week 48 of 2025, corresponding to the period from 29 September 2025 to 30 November 2025.

In both courses, the weekly instructional design followed a common structure. Each week, students were assigned an individual discussion topic, a group-based project submission, a personal weekly learning plan, and an end-of-week reflection that also served as the basis for planning the following week. In addition, lecture slides, instructional videos, and reference materials were uploaded to the LMS before or during the corresponding instructional week. The LMS course spaces, assignment submission functions, and activity logging mechanisms were prepared before the observation period began. Therefore, coursework submission records were available from the first instructional week for both courses.

In this study, an LMS event refers to a timestamped interaction record generated when a learner or instructor performs an action on the system. These events include, but are not limited to, logging into the LMS, accessing lecture materials, opening or viewing instructional videos, downloading reference materials, creating or updating a weekly learning plan, submitting an individual discussion post, commenting on peers' posts, uploading or resubmitting a group project, receiving or viewing grades and feedback, and submitting end-of-week reflections. Thus, the event data represent observed system interactions rather than merely planned instructional activities.

After preprocessing, the raw LMS events were aggregated into weekly learner-course indicators. Given 145 course-level student enrollments and nine instructional weeks, the maximum possible number of course-learner-week records was 1305 before preprocessing. The final analytical records were lower when students had incomplete activity sequences, withdrew from a course, or had no valid LMS records for a given week. Because the raw event logs were aggregated for analysis and later removed to reduce storage requirements, the exact total number of raw LMS activity records is no longer available. For this reason, this study reports the categories of raw LMS events and the retained weekly aggregated analytical structure rather than an exact raw event count.

Because the data were collected from regular LMS-supported teaching activities rather than from a controlled experimental setting, the dataset naturally reflected heterogeneous patterns of student participation. Some students actively engaged with the LMS through frequent access, assignment submissions, discussion activities, weekly planning, reflections, and peer interactions, whereas others showed limited or irregular interaction with the system. In addition, a small number of students withdrew from the courses during the observation period, resulting in incomplete learner-week records. These conditions produced sparse and uneven activity patterns across students, weeks, and feature groups. Rather than treating this variability as an artificial data issue, the study considers it a characteristic of naturalistic learning analytics data collected in real instructional settings. Nevertheless, this variability should be taken into account when interpreting the findings, particularly with respect to intervention evaluation and generalizability beyond the two programming courses examined in this study. Table 1 summarizes the main characteristics of the dataset used in this study.

Table 1. Main characteristics of the dataset used in this study.

Item	Description
Institution type	Higher education institution
Country/context	Vietnam/LMS-supported higher education context
Academic period	ISO weeks 40–48 of 2025
Calendar period	29 September 2025–30 November 2025
Course domain	Information technology/programming
Courses	Object-Oriented Programming; Mobile Programming
Number of courses	2
Number of students	145 course-level enrollments: 85 in Object-Oriented Programming and 60 in Mobile Programming
Number of instructors	2
Observation period	9 instructional weeks
Raw LMS activity records	Login/access events, content access, video/material views, discussion posts/comments, project submissions, grading/feedback events, weekly planning records, reflection records
Raw LMS event count	Not retained after weekly aggregation due to storage constraints
Final analytical records	Weekly aggregated learner-course observations derived from retained process-oriented indicators
Assignment submission availability	Coursework submission records were available from the first instructional week
Prediction horizon	Week-ahead prediction, from week t to week $t + 1$
Intervention period	Recommendations generated from week 44; follow-up intervention outcomes observed during ISO weeks 45–48
Outcome variable	Weekly combined performance score normalized to a 0–100 scale
Generalizability note	Single institution, two undergraduate programming courses

2.3. Weekly Aggregation and Observation Unit

To operationalize the proposed closed-loop learning analytics system in a temporally realistic manner, learning activity logs are aggregated at a weekly level. Each observation is defined by a composite key ($\text{course_id} \times \text{user_id} \times \text{iso_year} \times \text{iso_week}$), representing a unique learner–course–week instance. This unit of analysis is used because the study includes two course contexts, and a learner’s activity, assignments, feedback, and performance score are course-specific. The learner–course–week structure allows learning behaviors to be modeled as an evolving process within each course rather than as static snapshots or end-of-course summaries.

Weekly aggregation fulfills various methodological and pedagogical purposes. First, it aligns the analytics cycle with common instructional rhythms in higher education, where teaching activities, assignments, feedback, and learner self-regulation are generally organized every week. Structuring data at this level allows for near-real-time monitoring of learning processes and facilitates timely pedagogical decision-making and intervention planning, which are key objectives of early warning systems [18].

Secondly, organizing observations into consecutive weekly windows helps reduce the risk of temporal information leakage, which is a known issue in the offline evaluation of predictive models. This problem occurs when future information is accidentally used to inform past predictions. By limiting each weekly observation to data up to that week, this approach maintains temporal causality and supports more realistic model validation and deployment scenarios [19].

The learner–course–week representation offers a flexible basis for incorporating prediction, decision-making, intervention, and feedback within a closed-loop framework. Weekly snapshots enable the system to monitor short-term changes in learning behavior,

monitor subsequent learner responses after recommended support actions, and adjust decision-making policies in future cycles. This temporal structure is particularly suited for action-oriented learning analytics systems that seek to link process data, pedagogical decisions, and measurable educational outcomes [12].

2.4. Process-Oriented Learning Indicators

This study represents the learning process as a continuous and interactive activity by creating a set of process-oriented learning indicators. These indicators are derived from weekly aggregated activity logs on a learning management system (LMS). Instead of focusing on static learner attributes or end-of-term outcomes, the indicators aim to capture the evolving behaviors, participation patterns, and interactions of learners over time. The selection and arrangement of these indicators are grounded in established theoretical frameworks from educational psychology and learning analytics, ensuring that they are both pedagogically relevant and easy to interpret.

The indicators are organized into four complementary dimensions that reflect key aspects of the learning process in online and blended higher education environments: (i) self-regulated learning (SRL), (ii) engagement in assignments and assessments, (iii) social interaction and networking structures, and (iv) cognitive and behavioral engagement as characterized by the ICAP framework. Together, these dimensions provide a comprehensive view of learning behavior, facilitating short-term predictions, informed pedagogical decision-making, and effective intervention planning within the proposed closed-loop system.

Recent studies have further demonstrated that behavioral indicators derived from LMS activity logs, such as interaction patterns, task execution times, and interaction constructs, can effectively capture learning dynamics and support short-term performance prediction [20,21]. These process-oriented indicators provide a more actionable representation of learning behavior than static demographic traits, particularly in the context of early intervention systems.

2.4.1. Self-Regulated Learning (SRL) Indicators

Self-regulated learning indicators are designed to assess learners' ability to plan, monitor, and reflect on their learning activities. Based on established models of self-regulated learning, these indicators are organized into three main phases: planning, execution (or monitoring), and reflection [3,22,23].

In this study, self-regulated learning is operationalized through time-based and plan-oriented behavioral indicators derived from weekly LMS logs. Figure 2 illustrates how weekly self-regulated learning is represented and operationalized within the proposed system through planning, execution/monitoring, and reflection-related feedback indicators organized as a closed-loop structure. Rather than relying on self-reported SRL measures, the proposed indicators reflect learners' ability to set weekly study goals, regulate their learning time, and complete planned activities. Specifically, planning is captured through the declared weekly study target (*goal_minutes*), execution and monitoring are reflected by estimated actual study time and plan adherence metrics (*actual_minutes_est*, *adherence_pct*, *plan_on_time_rate*), and reflection is approximated through weekly plan completion (*final_completion_percent*). This representation aligns with cyclical SRL models while remaining grounded in observable LMS behaviors.

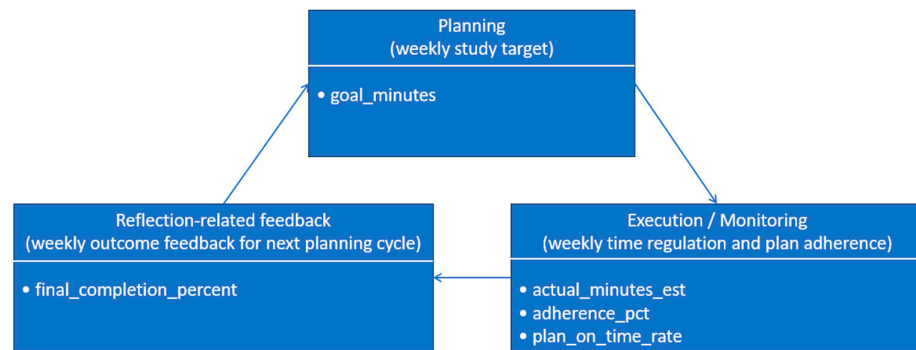


Figure 2. Weekly self-regulated learning (SRL) indicators and closed-loop structure based on planning, execution/monitoring, and reflection-related feedback derived from LMS activity logs.

2.4.2. Assignment and Assessment Indicators

Assignment and assessment indicators are designed to capture learners' engagement with formal learning tasks and evaluative processes over time. Unlike summative performance measures, these indicators focus on the behavioral and temporal aspects of task completion and feedback cycles. This focus is particularly useful for providing early warnings and formative instructional support.

In the proposed system, assignment and assessment engagement are measured through weekly aggregated LMS indicators that describe submission behavior, adherence to deadlines, and grading dynamics. Learners' effort and participation in assigned tasks are reflected in two key metrics: the number of submission attempts within a given week (*attempts_count*) and the number of tasks that have been graded (*graded_count*). These indicators offer a basic measure of engagement with assessment-related activities.

Timeliness is assessed through the on-time submission rate (*on_time_rate*) and the average delay in minutes for late submissions (*late_minutes_avg*). Together, these metrics characterize both compliance with deadlines and the severity of submission delays. Additionally, the dynamics of assessment feedback are represented by the average grading delay in days (*grade_delay_days_avg*), indicating how quickly evaluative feedback is made available to learners.

By organizing assignment and assessment indicators with a weekly temporal focus, the system treats assessment engagement as an ongoing process rather than a static outcome. Within the closed-loop learning analytics framework, these indicators support timely pedagogical decision-making by enabling the identification of burgeoning submission issues, feedback bottlenecks, or workload imbalances. Changes in these indicators following instructional interventions can be monitored in subsequent weeks, allowing the system to evaluate the effects of these interventions and iteratively refine decision-making policies.

2.4.3. Social Interaction and Network Structure Indicators

Social interaction and network structure indicators are intended to capture the participation patterns and roles of learners within an online learning community. In the context of learning analytics, these indicators reflect how learners engage with their peers through observable interaction behaviors and how those behaviors position them within the social structure of a course [17,24,25].

In the proposed system, social interaction is measured using weekly aggregated indicators derived from communication and discussion logs in the Learning Management System (LMS). The intensity of interaction is represented by the total number of interac-

tion events that occur within a given week (*interactions_total*), reflecting the overall level of participation among learners in social exchanges.

Additionally, network connectivity is characterized by the total degree (*degree_total*), which indicates how extensively a learner is connected to others within the interaction network. To describe learners' structural roles within the network, we use betweenness centrality (*betweenness*) as an indicator of intermediate or bridging positions in weekly interaction patterns. This measure reflects the extent to which a learner lies on interaction paths between peers, without implying any leadership or instructional authority.

Furthermore, relative influence in exchanges is captured by a normalized influence ratio (*influence_ratio_0100*), which indicates a learner's comparative prominence or reach within weekly interactions. The focus of interaction is illustrated by the reply concentration index (*reply_focus_0100*), showing whether learners' responses are distributed among many peers or directed toward specific interaction partners. Together, these indicators provide a process-oriented and comprehensible representation of social participation and network structure on a weekly basis.

Within the closed-loop learning analytics framework, changes in these social interaction indicators can signal emerging disengagement, social isolation, or shifts toward more central and reciprocal participation. Such signals complement self-regulation and assessment-related indicators, supporting timely and socially informed pedagogical decision-making and enabling the monitoring of intervention effects across subsequent weeks.

2.4.4. Engagement Indicators Based on the ICAP Framework

Engagement indicators are designed to characterize learners' observable patterns of participation and interaction during learning activities. In this study, engagement is defined using the ICAP framework, which distinguishes between Interactive, Constructive, Active, and Passive modes of engagement based on the observable actions of learners rather than inferred cognitive states [4].

In the proposed system, engagement based on the ICAP framework is implemented through weekly aggregated indicators that are derived from Learning Management System (LMS) discussion and activity logs. Basic participation intensity is measured using activity volume metrics, which include the number of original posts and comments (*posts*, *comments*), the total word count produced (*words_total*), and the average length of contributions per post (*avg_words_per_post*). These indicators reflect learners' levels of participation but do not assess the quality or effectiveness of their engagement.

To better understand the discourse-related aspects of engagement, the system includes indicators that reflect the structural properties of learners' written contributions. The average content depth (*depth_avg*) serves as a measure of the level of elaboration in weekly discussions. Additionally, the rate of comments that meet the Claim–Evidence–Reasoning (CER) structure (*cer_ok_rate*) indicates the presence of fundamental components of argumentation. These measures are combined to create a normalized argumentation index (*argumentation_index_0100*), which offers a clear representation of discourse organization on a larger scale.

Building on behavioral and discourse-related indicators, this study operationalizes ICAP engagement through normalized rates of constructive and interactive participation (*icap_constructive_rate_0100*, *icap_interactive_rate_0100*). These indicators reflect the relative proportion of higher-order engagement behaviors observed within a given week, based on learners' observable actions and contributions in the Learning Management System (LMS). Importantly, these rates represent behavioral manifestations of engagement rather than direct measures of learners' internal cognitive states. By emphasizing the prevalence of constructive and interactive activities, this approach aligns with the ICAP framework while avoiding assumptions about the underlying mental processes.

Lower-order engagement states, such as inactive or passive participation, are not modeled as separate ICAP indicators. Instead, these states are implicitly represented by the absence or low frequency of observable activities (e.g., zero posts, minimal interactions, or near-zero engagement rates). This design choice helps reduce redundancy among indicators and supports a concise and interpretable feature set that is suitable for short-term prediction and intervention.

By organizing ICAP-based engagement indicators on a weekly basis, the system captures short-term shifts in learners' participation and interaction profiles that are critical for early warning and pedagogical decision-making. Within the closed-loop learning analytics framework, changes in engagement profiles can signal emerging disengagement or positive responses to instructional support, complementing indicators related to self-regulated learning, assignments, and social networks.

2.4.5. Platform Usage and Time-on-Task Indicators

The system incorporates both theory-driven process indicators and additional metrics related to platform usage and task duration. These metrics help capture learners' overall engagement with the Learning Management System (LMS) environment by measuring the frequency and duration of their interactions without assuming specific learning strategies or engagement modes.

Platform usage is quantified through two primary measures: the total number of LMS sessions (*sessions_total*) and the total cumulative time spent on the platform per week (*dwell_total_min*). To provide a clearer understanding of how time is allocated across various learning activities, the dwell time is further divided into time spent on project-related work (*project_min*) and time dedicated to discussion-related activities (*discussion_min*). These indicators serve as essential behavioral signals that enhance self-regulated learning, assignment completion, social interaction, and overall engagement measures.

It is important to note that platform usage metrics should not be interpreted as direct indicators of learning quality. Instead, they offer contextual information about learners' presence and activity patterns during a specific week. This information can help identify early signs of disengagement and inform subsequent modeling and decision-making processes.

Table 2 summarizes the complete set of weekly process-oriented indicators used in this study, including self-regulated learning, assignment and assessment behaviors, social interaction and network measures, engagement and discourse indicators, and baseline platform usage metrics.

Table 2. Weekly process-oriented indicator taxonomy.

Indicator Group	Field (DB)	Brief Description	Unit/Scale
Self-Regulated Learning & Time Management	<i>goal_minutes</i>	Total planned study minutes per week	minutes
	<i>actual_minutes_est</i>	Estimated actual study minutes per week	minutes
	<i>plan_on_time_rate</i>	Plan adherence ratio (actual/goal, clipped to [0, 1])	0–1
	<i>adherence_pct</i>	Plan adherence ratio (actual/goal, clipped to [0, 1])	0–1
Assignments & Assessment	<i>attempts_count</i>	Number of assignment submission attempts per week	Count
	<i>graded_count</i>	Number of assignments graded during the week	Count
	<i>on_time_rate</i>	On-time submission rate	0–1
	<i>late_minutes_avg</i>	Average submission delay (late minutes)	Minutes
	<i>grade_delay_days_avg</i>	Average grading delay	Days
Social Interaction & Network (SNA)	<i>interactions_total</i>	Total interaction events (posts, replies, behavioral views)	Count
	<i>degree_total</i>	Total degree (network connectivity level)	Count
	<i>betweenness</i>	Betweenness centrality (bridging role in the network)	Index (>0)
	<i>influence_ratio_0100</i>	Influence ratio in exchanges (normalized 0–100)	0–100

	reply_focus_0100	Focused reply score (0–100)	0–100
	posts	Number of original posts	Count
	comments	Number of comments/replies	Count
	words_total	Total number of words written	Count
	avg_words_per_post	Average number of words per post	Count
	depth_avg	Average discourse depth	Index (>0)
Engagement/ICAP & Discourse	cer_ok_rate	Proportion of comments satisfying Claim–Evidence–Reasoning (CER)	0–1
	icap_constructive_rate_0100	Rate of constructive engagement (0–100)	0–100
	icap_interactive_rate_0100	Rate of interactive engagement (0–100)	0–100
	argumentation_index_0100	Argumentation quality index (0–100)	0–100
Platform Usage (LMS Baseline)	sessions_total	Number of LMS sessions	Count
	dwell_total_min	Total dwell time on the LMS	Minutes
	project_min	Time spent on projects or major assignments	Minutes
	discussion_min	Time spent on discussions/forums	Minutes
	prev_combined_score_0100	Previous week's combined performance score (0–100)	0–100

3. Predictive Modeling and Decision Framework

3.1. Prediction Target and Weakly Supervised Label Construction

The predictive objective of this study is to estimate short-term learning performance at a weekly granularity in order to support timely pedagogical decision-making within a closed-loop learning analytics system. Instead of predicting final course outcomes, the system focuses on week-ahead performance signals that are actionable during the learning process.

The prediction target is a continuous performance score defined on a standardized scale from 0 to 100. This score is constructed using a weakly supervised labeling strategy that integrates instructor assessment and peer feedback sentiment at the learner–course–week level. For learner i in course c during week t , the instructor score is computed from the graded coursework associated with that course-week and normalized to a 0–100 scale:

$$S_{i,c,t}^{instr} = \frac{1}{|G_{i,c,t}|} \sum_{g \in G_{i,c,t}} 100 \times \frac{score_g}{maxscore_g} \quad (1)$$

where $G_{i,c,t}$ is the set of graded assignments or project-related assessment records of learner i in course c during week t . $score_g$ denotes the awarded score for graded item g , and $maxscore_g$ denotes the maximum possible score for that item. When the original grade was already recorded on a 0–100 scale, the normalized value was used directly.

In parallel, peer feedback is incorporated through sentiment analysis of classmates' textual comments on submitted assignments or discussion-related learning outputs. The peer sentiment score is defined as the proportion of positive or supportive comments ($N_{i,c,t}^+$) relative to the total number of valid peer feedback comments ($N_{i,c,t}$) received by learner i in course c during week t , mapped to the same 0–100 scale:

$$S_{i,c,t}^{peer} = \frac{N_{i,c,t}^+}{N_{i,c,t}} \times 100 \quad (2)$$

The final weekly label is obtained by combining the normalized instructor score and the peer sentiment score using a fixed weighted aggregation function:

$$y_{i,c,t} = \alpha S_{i,c,t}^{instr} + (1 - \alpha) S_{i,c,t}^{peer} \quad (3)$$

where α controls the relative contribution of instructor assessment and peer feedback sentiment. In the current implementation, α was set to 0.8 giving instructor assessment

an 80% contribution and peer sentiment a 20% contribution when valid peer feedback was available. When peer feedback was unavailable, the label relied solely on the instructor assessment component:

$$y_{i,c,t} = S_{i,c,t}^{instr} \quad (4)$$

Importantly, sentiment information is used exclusively for label construction and is not included among the predictive features. This design reduces the risk of information leakage and preserves a clear separation between observed learning behaviors used as model inputs and the supervision signal used for training. Because the target combines formal instructor assessment with a supplementary formative signal from peer feedback, it should be interpreted as a weakly supervised weekly performance label rather than as a direct measure of final course achievement. By adopting a continuous performance score rather than binary risk labels, the proposed target supports flexible ranking- and threshold-based intervention strategies within the decision-oriented closed-loop learning analytics system.

3.2. Predictive Modeling Approach

3.2.1. Prediction Formulation and Input Features

In this study, the prediction task is framed as a supervised regression problem. The goal is to estimate the weekly performance score $y_{i,c,t+1}$ for learner i in course c in the following week $t + 1$, based on a set of process-oriented indicators observed up to week t . Formally, the model aims to learn a relationship between weekly learner-course behavioral features and a continuous outcome variable defined on a scale from 0 to 100.

Given the diverse nature of the input features—including time-based indicators, proportion-based measures, network statistics, and discourse-related indices—the modeling approach focuses on algorithms capable of handling mixed feature types, non-linear relationships, and interaction effects, all while remaining resilient to the noise commonly found in learning management system logs.

3.2.2. Model Selection and Rationale

In the current implementation, the predictive component was evaluated using several regression models, including a previous-week score baseline, linear regression, random forest regression, and LightGBM regression. The baseline and comparative models were included to examine whether more complex models provide added value beyond simpler prediction approaches and recent performance history.

Tree-based ensemble models were considered relevant because they can capture non-linear relationships and interaction effects among heterogeneous LMS indicators. Gradient boosting constructs an ensemble of decision trees in a sequential manner, where each successive tree is trained to reduce the residual errors of the previous ensemble. This iterative refinement enables gradient boosting models to capture complex, non-linear relationships and interaction effects among heterogeneous input features, which are common in learning management system data [26–28].

Compared to linear models, tree-based ensemble methods are more flexible in modeling non-additive relationships between behavioral indicators, such as the interplay between self-regulated learning, assessment behavior, platform usage, and social interaction [26]. In addition, tree-based ensemble methods are relatively insensitive to feature scaling and can accommodate mixed feature types, missing values, and skewed distributions [29,30]. These properties make gradient boosting particularly well-suited for naturalistic educational data derived from learning management systems, where behavioral signals are heterogeneous and often incomplete [28].

In recent years, alternative approaches based on deep learning have also been proposed for student performance prediction, including recurrent neural networks (RNN), long short-term memory (LSTM) models, and transformer-based architectures that capture temporal dependencies in learning behaviors [31–33]. While these models demonstrate strong performance in sequence modeling tasks, they often require larger-scale datasets, involve higher computational complexity, and present additional challenges in interpretability, particularly in educational contexts where transparency and instructor trust are important.

In the current deployment context, the choice of model is therefore interpreted as a trade-off among predictive performance, robustness, interpretability, and operational usefulness. The comparative evaluation showed that linear regression achieved the strongest predictive performance in the current dataset, while LightGBM was retained for the subsequent explainability and intervention-planning workflow because of its stable predictive performance and practical compatibility with SHAP-based feature attribution. This choice does not imply that LightGBM is the most accurate model in the dataset; rather, it reflects the need to connect prediction outputs with local explanations and instructor-reviewable recommendations.

3.2.3. Baseline Models and Robustness Checks

To evaluate whether the proposed predictive model provides added value beyond simpler approaches, several baseline models were included in the evaluation. First, a previous-week score baseline was used, in which the predicted score for week ($t + 1$) was set equal to the observed performance score from week (t). This baseline is important because short-term learning performance often exhibits temporal continuity, and recent performance may already provide a strong prediction signal.

Second, a linear regression model was used as a simple parametric baseline to assess whether the relationship between weekly learning indicators and performance could be adequately captured by an additive linear model. Third, a random forest regression model was included as a non-linear ensemble baseline that does not rely on sequential boosting. Finally, the LightGBM-based model was evaluated as the main gradient boosting approach.

Model performance was assessed using mean absolute error (MAE), root mean squared error (RMSE), coefficient of determination (R^2), and Spearman rank correlation. MAE and RMSE measure absolute prediction error, while Spearman correlation evaluates whether the model preserves the relative ordering of students, which is important for rank-based intervention selection. In addition, an ablation check was conducted by removing the previous-week performance feature (`prev_combined_score_0100`) to examine the extent to which prediction accuracy depended on recent historical performance.

These baseline and robustness checks were included to address model complexity and to determine whether the proposed model provides practical predictive benefit beyond simpler alternatives and temporal continuity.

3.2.4. Time-Aware Training and Validation Protocol

To ensure temporal validity and realistic deployment conditions, model training follows a strictly time-aware protocol. For each prediction cycle corresponding to week ($t + 1$), the model is trained using learner–course observations aggregated from weeks 1 through (t). This expanding-window strategy reflects actual operational usage, where predictions depend solely on historical observations, and prevents the leakage of future data.

Validation is conducted using forward-chaining splits that align with academic weeks, allowing performance to be assessed under conditions that mirror real-time system operation. The dataset is ordered chronologically by ISO year and ISO week, and

model evaluation is performed on later weeks that are not used for training. This design provides a more realistic estimate of week-ahead prediction performance than random shuffling, which could mix future and past observations.

All input features are standardized or normalized as necessary to ensure numerical stability. Missing values resulting from sparse activity or unavailable weekly records are handled through the internal mechanisms of the model or through consistent preprocessing rules defined before model training. This design preserves the natural structure of LMS data and avoids introducing artificial signals into the learning process.

3.2.5. Explainability Integration

An additional motivation for selecting gradient boosting models is their compatibility with post hoc explainability techniques, particularly SHAP (SHapley Additive exPlanations). SHAP provides a theoretically grounded framework for attributing a model's prediction to individual feature contributions based on cooperative game theory. In this study, SHAP values are used to explain both global patterns across the student population and local, instance-level predictions at the learner-course-week level [34,35].

This explainability capability is important in a closed-loop learning analytics system, where predictive outputs must be interpretable and reviewable by instructors and academic support staff. By combining predictive modeling with transparent feature attribution, the system supports short-term performance estimation while also providing a basis for understanding which observable behavioral indicators are associated with a learner's predicted performance.

However, SHAP explanations are interpreted as model-based feature attributions rather than causal effects. A negative SHAP value indicates that a feature contributes to lowering the model's predicted score for a particular learner-course-week instance, but it does not prove that directly changing that feature will necessarily cause an improvement in performance. This distinction is important because the subsequent intervention recommendations should be understood as pedagogically motivated and model-informed suggestions rather than causal prescriptions.

3.2.6. Prediction Output for Decision-Making

The model produces a continuous prediction $\hat{y}_{i,c,t+1} \in [0, 100]$, representing the estimated short-term performance level of learner i in course (c) in the upcoming week. This score-based output is intentionally retained in its continuous form and is not thresholded at the modeling stage. Maintaining a continuous prediction allows downstream decision policies to flexibly adapt intervention strategies based on ranking, utility optimization, or resource constraints, as described in Section 3.3.

By integrating temporally constrained training protocols, baseline comparisons, robustness checks, explainability integration, and continuous score prediction, this approach supports interpretable short-term performance estimation while remaining closely linked to the system's closed-loop decision framework. The inclusion of baseline and ablation analyses also helps clarify whether the proposed model offers added predictive value beyond simpler approaches and recent performance history.

3.3. Decision Framework: From Scores to Pedagogical Actions

While accurate predictions are useful for an early warning system, they alone do not adequately support decision-making in educational practices. In a closed-loop learning analytics system, predicted outcomes need to be transformed into specific support decisions that consider available resources, instructional constraints, and instructor review. Therefore, this study utilizes a decision-oriented theoretical framework, clearly distinguishing between prediction and action, and emphasizes how to effectively apply contin-

uous performance scores in the context of teaching. Integrating predictive analytics with decision-making processes is increasingly recognized as a crucial direction in learning analytics research. Recent studies highlight the need for decision support systems that not only identify learners at risk but also optimize intervention strategies under resource constraints [36,37]. By clearly separating prediction from decision-making policies, this proposed framework aligns with emerging approaches in cognitive decision-making analytics, where model outputs are evaluated based on practical usefulness rather than predictive accuracy alone.

At each weekly cycle, the predictive model outputs a continuous performance score $\hat{y}_{i,c,t+1} \in [0, 100]$ for each learner i in course c , representing the estimated short-term performance level in the upcoming week. Lower scores indicate a higher likelihood of short-term learning difficulties. This continuous representation preserves relative differences between learners and avoids premature discretization into binary risk labels.

The first decision mechanism is a rank-based selection strategy. Learners are sorted in ascending order according to their predicted scores, and the bottom K learners or a fixed proportion of the cohort are selected as candidates for instructor review and possible support. The value of K is determined by available instructional or support resources, such as advising capacity, tutoring availability, or the number of students an instructor can reasonably follow up within a week.

This strategy requires no predefined performance threshold and adapts naturally to cohort size and score distribution. As a result, it is stable under week-to-week performance fluctuations and well-suited to operational settings with fixed support capacity.

To support contexts where intervention costs and benefits must be explicitly balanced, the system also implements a threshold-based decision policy. An intervention threshold τ is defined such that learners with $\hat{y}_{i,c,t+1} \leq \tau$ are selected as candidates for review and support.

The threshold can be selected by considering an intervention utility function that accounts for the relative benefits of correctly identifying learners in need and the costs associated with unnecessary or missed support. Threshold selection is guided by Decision Curve Analysis (DCA), which evaluates the net benefit of different decision thresholds across a range of pedagogically meaningful trade-offs [38].

Decision policies are reviewed on a weekly basis. After recommendations are reviewed and support actions are implemented where appropriate, subsequent learner behaviors and outcomes are observed in the LMS and incorporated into the next prediction cycle. This feedback enables instructors to adjust ranking cutoffs or intervention thresholds in response to changing cohort dynamics, instructional constraints, and observed learner responses. However, these observations should be interpreted as preliminary feedback rather than causal evidence of intervention effectiveness.

By decoupling prediction, decision policy, instructor review, and intervention planning, the proposed framework supports flexible and transparent pedagogical decision-making while remaining tightly integrated with the closed-loop learning analytics architecture.

In the current system deployment, the rank-based selection strategy is adopted as the primary decision mechanism. This choice is motivated by its simplicity, robustness, and direct alignment with practical instructional constraints, where support capacity is typically fixed and determined in advance. Rank-based selection enables stable weekly planning without requiring explicit specification of intervention costs, which are often difficult to quantify reliably in real educational settings.

3.4. From Explanation to Action: XAI-Guided Intervention Planning Design

Following the score-to-action analysis framework in Section 3.3, the next step is to translate predictions into feasible intervention plans at the learner-week level. To achieve this, the system integrates a two-level Explainable Artificial Intelligence (XAI) layer, linking model predictions to explainable behavioral factors and feasible pedagogical actions.

At the general level, global SHAP analysis is used to estimate the relative importance of trait groups, including self-regulating learning, homework behavior, social interaction, engagement/discussion (ICAP), and recent activity history. These aggregate explanations provide an overview of the dominant behavioral drivers in each course or semester, supporting guidance for adjusting strategies at the learner-course-week level (e.g., strengthening on-time submission policies or enhancing structured interaction). Recent research has emphasized the importance of explainable artificial intelligence (XAI) in educational contexts, where transparency and interpretability are essential for building trust and supporting pedagogical decision-making [39–41]. In learning analytics, XAI techniques enable instructors to understand the underlying behavioral factors influencing student performance and to design targeted interventions accordingly. This is particularly important in process-oriented systems, where actionable insights must be derived from complex behavioral data rather than static attributes.

Figure 3 presents the global SHAP importance aggregated at the feature-group level for week ISO 2025–W44 [42]. The results indicate that platform-related activity, assignment and assessment behavior, and recent learning history are the three most influential groups, with mean SHAP values of 2.30, 2.27, and 2.18, respectively. These groups substantially outweigh other behavioral dimensions in terms of their contribution to short-term performance prediction. The prominence of platform- and assignment-related metrics suggests that immediately observable learning behaviors—such as system time, submission patterns, and assessment engagement levels—play a dominant role in determining weekly performance. Similarly, the strong influence of recent historical traits highlights the temporal continuity of learning behavior, where recent performance and operational patterns serve as reliable predictors of near-future outcomes. In contrast, ICAP-based discourse indicators and self-regulated learning (SRL) measures exhibit relatively lower importance in this setting, while social interaction (SNA) features contribute minimally to the prediction. This does not imply that these dimensions are unimportant for learning; rather, their effects may be more indirect, delayed, or context-dependent, and therefore less detectable in short-term, week-ahead prediction tasks. Overall, the global explanation reveals that short-term performance in this context is primarily driven by immediate behavioral signals and recent activity patterns, which aligns with the design of the weekly prediction framework and supports the use of actionable, time-sensitive intervention strategies.

At the individual level, for each learner i at week t , the predicted score $\hat{y}_{i,t+1}$ is decomposed into feature contributions using SHAP. Features with large negative SHAP values are identified as dominant risk factors that reduce predicted performance. These factors are mapped to interpretable behavioral categories, enabling instructors to understand the underlying causes of risk at the learner-week level.

Table 3 illustrates typical local explanations for the three learners, highlighting the prominent characteristics contributing to the reduction in predicted performance in week $t + 1$. For each learner, the characteristic with the largest negative SHAP value was identified as the primary risk factor. For St-A students, the most influential factor was `project_min` (platform-related activity), with a SHAP value of -1.52 , indicating insufficient time spent on project work significantly reduced the predicted score. Conversely, for St-B and St-C students, the primary factor was `on_time_rate`, belonging to the assignment and assessment group. Negative SHAP values (-0.88 and -0.25 , respectively) indicate that late

or inconsistent submission behavior contributed to performance risk, with a stronger impact observed in St-B.

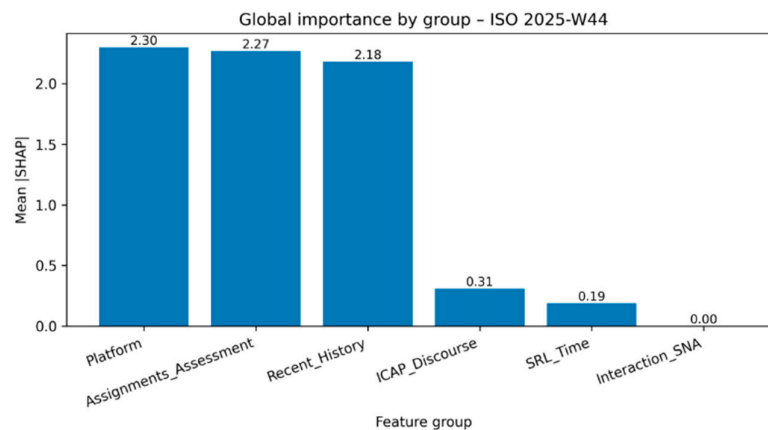


Figure 3. Global SHAP importance aggregated by feature groups for week ISO 2025–W44. Values represent mean absolute SHAP contributions across all learners.

Table 3. Examples of local SHAP explanations for selected learners at week ISO 2025–W44. The table reports the most influential negative feature for each learner, its corresponding feature group, SHAP value, and predicted performance score.

Student	ISO Year	ISO Week	Objects	Attribute	Group Attribute	Shap Value	Prediction Point
St-A	2025	44	Mobile App programming	project_min	Platform	−1.52	53.24
St-B	2025	44	Mobile App programming	Ontime_rate	Assignments Assessment	−0.88	62.86
St-C	2025	44	Mobile App programming	Ontime_rate	Assignments Assessment	−0.25	81.94

These examples demonstrate how local SHAP interpretations provide personalized diagnostic insights by linking predicted outcomes to specific behavioral deficits. Furthermore, the differences in salient traits among practitioners confirm that performance risk is not caused by a single factor but stems from diverse behavioral patterns, further reinforcing the need for personalized intervention strategies.

Based on the explanation layer, we construct counterfactual recommendations in the form of a 2–3 step plan, governed by two key components.

The first component defines an intervention mapping space rather than a direct action space for manipulating behavioral variables. In this study, process-oriented indicators such as study time (actual_minutes_est), submission timeliness (on_time_rate), average lateness (late_minutes_avg), and engagement-related measures such as practice attempts (attempts_count) are treated as modifiable behavioral targets or expected short-term outcomes. The actual interventions are pedagogical actions that instructors can review and implement, such as sending deadline reminders, requiring interim progress checkpoints, providing weekly study-planning guidance, offering structured discussion prompts, or recommending additional practice activities.

This distinction is important because the system does not directly change student behavior or causally manipulate LMS indicators. Instead, it maps risk-related indicators to feasible instructor-mediated actions that are intended to support improvement in the corresponding behavioral targets during the following instructional week. In contrast,

static attributes or variables that cannot be reasonably influenced through short-term pedagogical support, such as identity-related features or structurally fixed characteristics, are excluded from the intervention mapping space.

The second component specifies feasibility constraints to ensure that recommended actions remain realistic and executable. First, the magnitude of change is restricted to small and realistic ranges (e.g., increasing study time by approximately 60 min, adding two constructive comments, completing one additional on-time submission, or reducing lateness by around 10 min). Second, all adjustments must respect valid technical boundaries, such as maintaining rates within the interval $[0, 1]$ and ensuring non-negative values for time-related variables. Third, the complexity of intervention plans is limited to a maximum of three actions per learner, in order to maintain clarity and ease of execution within a weekly instructional cycle.

Intervention plans are generated by using a counterfactual recourse approach. Candidate actions are applied individually or in small combinations to modify selected features, and the predictive model is re-evaluated to estimate the resulting change in predicted performance:

$$\Delta\hat{y} = \hat{y}_{i,t+1} - \hat{y}_{i,t} \quad (5)$$

The selected intervention plan maximizes expected improvement $\Delta\hat{y}$ while minimizing total action cost. This process ensures that recommended actions are both effective and feasible within the weekly instructional cycle.

To operationalize the previously defined action space $\mathcal{A}$ for intervention planning, each action $a \in \mathcal{A}$ is parameterized by a set of attributes, including a target variable (feature), unit of measurement, permissible value range, minimum adjustment step, and a normalized cost $c(a)$ defined on a scale of 1 to 1.5. This parameterization enables the formulation of cost-aware intervention plans under practical constraints, including a maximum of three actions per learner and a total cost constraint $\sum c(a) \leq 3$. Table 4. A feasible action space is defined for weekly intervention planning. Each action specifies a modifiable behavioral feature, its allowable range, adjustment step, and normalized implementation cost.

Table 4. Mapping from risk indicators to pedagogical intervention actions and expected behavioral targets.

Code	Risk Indicator/ Target Feature	Instructor-Facing Intervention	Student-Facing Action	Expected Behavioral Target	Feasibility Constraint	Cost
A01	Low actual_minutes_est	Provide weekly study-time planning and progress monitoring	Allocate additional scheduled study time	Increase estimated study time	+60 to +120 min	1.0
A02	Low on_time_rate	Send deadline reminder and require an interim checkpoint	Submit draft or progress checkpoint before deadline	Improve on-time submission behavior	+49 to +99 percentage points	1.5
A03	Low icap_constructive_rate_0100	Provide structured discussion prompt or evidence-based response scaffold	Submit one constructive discussion response	Increase constructive participation	+10 point	1.2
A04	Low interactions_total	Assign small-group peer interaction task	Reply to or interact with at least three peers	Increase meaningful interactions	+3 interactions	1.0
A05	Low depth_avg	Provide response template or CER writing scaffold	Revise response with claim, evidence	Improve depth of written response	+0.3 score	1.3

			dence, and reasoning			
A06	Low attempts_count	Recommend additional formative practice or quiz attempt	Complete one additional practice attempt	Increase practice attempts	+1 attempt	1.0
A07	Hight late_minutes_avg	Provide time-management reminder and early submission plan	Plan completion before the deadline	Reduce average lateness	−60 min	1.0
A08	Low dwell_total_min	Recommend targeted review of learning materials	Review lecture video or reference material	Increase time on learning materials	+60 min	1.0

Table 4 presents the intervention mapping space defined for weekly intervention planning in the proposed system. Each row links a risk-related LMS indicator to an instructor-facing pedagogical intervention, a student-facing action, an expected behavioral target, a feasible adjustment range, and a normalized implementation cost. This structure distinguishes between the observed indicator that signals risk, the pedagogical action that can be implemented by the instructor, and the behavioral outcome that the intervention is intended to support.

The recommended interventions focus on short-term learning behaviors that can be supported within a one-week instructional cycle, including study-time planning, submission timeliness, engagement intensity, peer interaction, and discourse quality. For example, improving on-time submission rate or reducing average lateness is treated as an expected behavioral target rather than the intervention itself. The corresponding pedagogical interventions may include deadline reminders, interim progress checkpoints, weekly study-planning guidance, or submission checklists. In contrast, non-actionable or static attributes are excluded to ensure that the recommendations remain feasible and pedagogically appropriate.

The inclusion of feasible adjustment ranges and normalized implementation costs ensures that candidate recommendations remain within valid and interpretable bounds. The adjustment ranges represent small short-term behavioral targets that may reasonably be pursued during the following instructional week, while the normalized cost reflects the relative instructional effort required to implement the recommendation. These costs and adjustment ranges are expert-defined operational parameters in the current implementation and should not be interpreted as empirically estimated causal effects.

In deployment, the XAI-to-recourse layer provides three practical advantages. First, it enhances transparency by enabling instructors to understand why a learner is flagged through feature-level explanations grounded in observable LMS behavior. Second, it supports practical feasibility by generating small, concrete, and instructor-reviewable recommendations that can be implemented within a short instructional cycle. Third, it supports subsequent monitoring by allowing predicted improvements to be compared with observed outcomes in later weeks. However, such comparisons are used only as preliminary feedback for refining the system and should not be interpreted as causal evidence that the recommended actions improved learner performance without further controlled or quasi-experimental validation.

3.5. Operational Cycle and Impact Measurement (Closing the Loop)

The proposed system is designed as a closed-loop learning analytics framework that operates on a weekly cycle, integrating prediction, decision support, recommendation generation, and feedback logging into a unified process. At each week t , learner interaction data are collected from the LMS and transformed into process-oriented features as

described in Section 2. These features are used to generate predicted performance scores for week $t + 1$, which are subsequently translated into support decisions and actionable recommendations through the decision and XAI-to-recourse layers.

Once recommendations are reviewed and implemented where appropriate, learners' subsequent behaviors and outcomes are observed during week $t + 1$ and recorded in the LMS. These observations serve two purposes. First, they provide updated input features for the next prediction cycle, enabling the model to continuously adapt to evolving behavioral patterns. Second, they constitute preliminary feedback of learner response to interventions, forming the basis for impact assessment within the closed-loop framework.

To support preliminary monitoring, the system is designed to compare the predicted improvement suggested by recourse actions ($\Delta\hat{y}$) with the observed change in actual performance (Δy) in subsequent weeks. This comparison provides a feedback signal for assessing the consistency and reliability of intervention recommendations. In particular, alignment between $\Delta\hat{y}$ and Δy may indicate that the recommendations are directionally consistent, while systematic discrepancies may reveal overestimation, underestimation, or context-specific limitations of the intervention planning design.

Although this study focuses on system design and preliminary deployment rather than controlled evaluation of intervention effects, the proposed framework enables systematic feedback logging over time. The recorded feedback can be used to refine recommendation selection strategies, adjust ranking cutoffs or decision thresholds, and improve the correspondence between predicted and observed learning dynamics without modifying the core system architecture. However, such feedback should not be interpreted as causal evidence of intervention effectiveness without further controlled or quasi-experimental validation.

All predictions, decision outputs, recommended actions, and subsequent outcomes are logged in a structured audit trail. This audit mechanism supports transparency and reproducibility in both prediction and recommendation processes. It can also support future longitudinal analyses of system performance, including utility estimation, calibration of decision policies, and updates to decision curves across academic periods.

Through this iterative process, the system moves beyond static prediction toward feedback-informed decision support, where predictive models, decision policies, and recommendation mappings can be refined based on observed outcomes. This design operationalizes the concept of closed-loop learning analytics by explicitly linking prediction, instructor-reviewable action, and feedback logging within a continuous cycle of improvement.

4. Results and Discussion

This section presents the experimental results of the proposed closed-loop learning analytics system in a real-world teaching environment. The evaluation focuses not only on the predictive performance of the model, but also on how these predictions translate into feasible intervention strategies and their practical implications for teaching and learning. With the system's two main objectives—accurate short-term performance estimation and decision-oriented pedagogical support—the results are organized into two parts. The first examines the predictive behavior and generalizability of the model, while the second analyzes the effectiveness and limitations of the intervention mechanisms derived from the model's output.

4.1. Predictive Performance and Model Behavior

The first part of the analysis focuses on the predictive performance and model behavior of the proposed system. The evaluation considers not only absolute prediction accuracy, but also ranking performance and robustness across different model configura-

tions. This is important because the proposed decision framework uses continuous prediction scores to support rank-based and resource-sensitive intervention selection.

To reflect realistic forecasting conditions, the dataset was ordered chronologically based on ISO year and ISO week. The final 20% of the learner–course–week observations were reserved as the independent test set, while the remaining observations were used for training and validation without random shuffling. This time-based split ensures that later observations are predicted from earlier records and reduces the risk of temporal information leakage.

Table 5 presents the comparative predictive performance of the baseline and regression models on the test set. Among the evaluated models, linear regression achieved the strongest overall predictive performance, with the lowest MAE value (3.5573), lowest RMSE value (4.6721), highest R^2 value (0.7073), and highest Spearman rank correlation ρ value (0.8399). This finding suggests that, for the current dataset and time-based evaluation setting, the relationship between weekly LMS-based indicators and short-term performance can be captured effectively by a relatively simple additive model.

Table 5. Comparative predictive performance of baseline and regression models on the test set.

Model	MAE	RMSE	R^2	Spearman ρ
Previous-week score baseline	3.9490	5.3016	0.6232	0.7964
Linear regression	3.5573	4.6721	0.7073	0.8399
Random forest regression	3.9500	5.2813	0.6260	0.8020
LightGBM regression	3.8863	5.1700	0.6416	0.8025
LightGBM without previous score	4.9177	6.0545	0.5085	0.7061

The previous-week score baseline also performed competitively, with MAE = 3.9490, RMSE = 5.3016, R^2 = 0.6232, and Spearman ρ = 0.7964. This confirms that recent performance history provides an important signal for week-ahead prediction. Random forest regression and LightGBM regression showed comparable performance, but neither outperformed linear regression in this evaluation.

The LightGBM model achieved MAE = 3.8863, RMSE = 5.1700, R^2 = 0.6416, and Spearman ρ = 0.8025. Although LightGBM did not achieve the best predictive performance, it was retained for the subsequent explainability and intervention-planning workflow because it provides stable prediction performance and supports SHAP-based feature attribution. Therefore, LightGBM should be interpreted as the model used for the explanation-to-action demonstration rather than as the most accurate model in the current dataset.

The ablation result further highlights the importance of the previous-week performance feature. When (prev_combined_score_0100) was removed from the LightGBM model, performance declined substantially, with MAE increasing to 4.9177, RMSE increasing to 6.0545, R^2 decreasing to 0.5085, and Spearman correlation decreasing to 0.7061. This indicates that recent performance history plays a central role in week-ahead prediction.

Figure 4 illustrates the relationship between predicted and actual scores for the LightGBM model used in the explainability and intervention-planning workflow. The predicted values generally follow the diagonal trend, indicating that the model captures the overall association between process-oriented LMS indicators and short-term performance outcomes. At the same time, the model shows some dispersion around the diagonal, suggesting that individual-level prediction remains uncertain, particularly for learners with more extreme performance patterns.

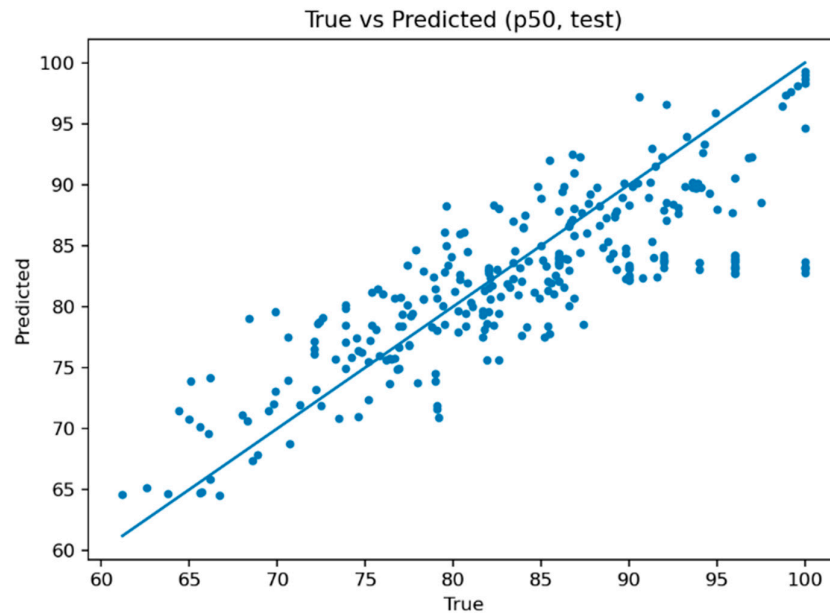


Figure 4. True vs. predicted performance scores on the test set for the LightGBM model used in the explainability workflow.

To assess the training behavior of the LightGBM model used for the explainability workflow, the learning curve was examined. As shown in Figure 5, the validation MAE decreased rapidly during the early boosting iterations and then stabilized at approximately 3.9 on the normalized 0–100 scale. In contrast, the training MAE continued to decrease as the number of estimators increased. This pattern suggests that the model captured the dominant predictive signal relatively early, while additional boosting iterations mainly improved the fit to the training data. Therefore, the learning curve supports the use of LightGBM as a stable model for SHAP-based explanation, but it should not be interpreted as evidence that LightGBM outperforms the simpler baseline models.

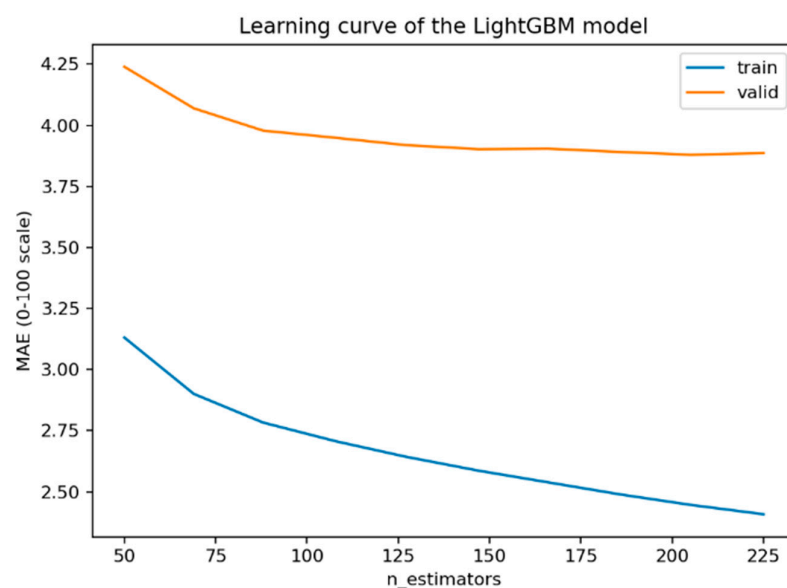


Figure 5. Learning curve of the LightGBM model used in the explainability workflow, showing training and validation MAE across boosting iterations.

Overall, the comparative results show that the predictive task in this dataset can be addressed effectively by relatively simple models. Linear regression achieved the strongest predictive performance, while the previous-week score baseline also performed competitively, indicating strong temporal continuity in weekly learning performance. LightGBM provided stable prediction performance and was retained for SHAP-based explanation and intervention planning, but it should not be interpreted as the most accurate model in the current evaluation.

These findings help moderate the interpretation of the predictive component. The contribution of the proposed system lies not in achieving superior prediction accuracy over all baselines, but in operationally connecting temporally valid prediction, explainable AI, instructor-reviewable intervention planning, and feedback logging within a closed-loop learning analytics workflow.

4.2. Intervention-Planning Outputs and Practical Implications

Beyond predictive accuracy, the core purpose of the proposed system is to translate model outputs into instructor-reviewable pedagogical recommendations. This subsection illustrates how the proposed system generates intervention plans from predicted scores, SHAP-based explanations, and predefined recommendation mappings. The selected learners are representative examples used to demonstrate the explanation-to-action workflow rather than to provide statistical evidence of intervention effectiveness. Therefore, the following analysis should be interpreted as a preliminary observational illustration of the recommendation workflow, not as causal evidence that the suggested actions improved learner performance.

The proposed system was deployed in a real instructional setting and operated on a weekly basis. During this period, learner interaction data were collected weekly from the LMS, processed into behavioral indicators, and used to generate short-term performance predictions, ranking information, SHAP-based explanations, and instructor-reviewable recommendations.

For evaluation purposes, we focus on a four-week window from ISO week 44 to week 47, during which the system outputs were consistently available. This period allows us to examine the relationship between predicted scores and subsequent observed performance trajectories. At each weekly cycle, the system produces a structured set of outputs for each learner, including: (i) a predicted performance score, (ii) a ranking position for candidate selection, (iii) dominant risk factors identified through SHAP-based local explanations, and (iv) a short-term recommendation plan derived from the intervention mapping space described in Section 3.4.

Recommendation design follows the framework described in Section 3. Behavioral indicators are analyzed first at the overall level to understand general patterns, and then at the individual level to identify learner-specific risk factors. Based on these diagnoses, instructor-facing recommendations are selected from the predefined intervention mapping space under feasibility and cost constraints. Table 6 presents representative recommendation plans generated by the system for selected learners during week 44. These examples illustrate how risk factors are mapped to instructor-reviewable support actions and expected behavioral targets.

Table 6. Illustrative recommendation plans derived from SHAP-based explanations and their expected model-based improvement in week 44. The selected learners are representative examples used to demonstrate the explanation-to-action workflow rather than to estimate intervention effectiveness.

Student	ISO Week	Dominant Risk Factor	Instructor-Facing Recommendation	Expected Behavioral Target	$y_{i,t}$	$\hat{y}_{i,t+1}$	Δ (Change)
St-A	44	Low project_min/low study time	Provide weekly study-time planning and progress monitoring	Increase estimated study/project time by +120 min	53.24	54.55	1.31
St-B	44	Low on_time_rate	Send deadline reminder and require interim checkpoint	Improve on-time submission behavior by approximately +49 percentage points	62.86	64.57	1.71
St-C	44	Low on_time_rate; limited expected gain near upper range	Provide planning reminder and submission checklist	Improve on-time submission behavior/maintain timely completion	81.94	83.12	1.18

Table 6 presents representative intervention plans generated by the system for selected learners at week 44. For each learner, the table reports the predicted score before applying the recommendation plan, the predicted score after applying the model-based recourse plan, and the corresponding expected change $\Delta\hat{y}$. The recommendations are derived from dominant SHAP-based risk factors and mapped to feasible instructor-facing support actions.

The expected improvements range from 1.18 to 1.71 points. These values should be interpreted as model-based expected changes rather than observed causal effects. The results indicate that the recourse mechanism can identify feasible recommendation plans that are expected to increase the model-predicted score under the model assumptions. However, whether these recommendations lead to actual performance improvement depends on learner response, implementation quality, and contextual factors that are not fully captured by the model.

While Table 6 summarizes the recommendation plans and their expected effects, Figure 6 provides a temporal view of predicted and observed performance over consecutive weeks. This visualization allows us to examine the alignment between predicted trajectories and subsequent observed outcomes, but it does not by itself establish causal intervention effects.

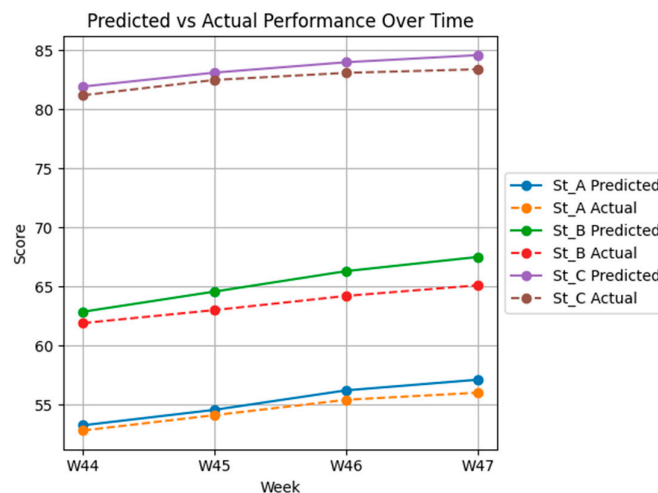


Figure 6. Illustrative predicted and observed performance trajectories over four consecutive weeks (W44–W47) for three representative learners. Solid lines denote predicted scores, while dashed lines

denote observed scores. The figure is intended to support preliminary feedback monitoring and should not be interpreted as causal evidence of intervention effects.

For St_A, the predicted and actual curves nearly coincided throughout the period, with only small gaps between them. This suggests that the predicted and observed trajectories were directionally aligned for this learner, and that the model provided a reasonable short-term estimate of the observed performance trend. For St_B, a consistent gap was observed between predicted and actual performance, with the predicted score systematically higher. This suggests a tendency for the model to overestimate the expected improvement, possibly due to incomplete implementation of the recommended actions or external learning constraints. For St_C, both predicted and actual scores remained relatively high, with a slower rate of improvement over time. The slower increase in the curves suggests a possible diminishing-return pattern, where gaining further benefit becomes more difficult as performance approaches the upper range. In this case, the model accurately captured the trend but slightly overestimated the incremental improvement.

The gap between the predicted and observed trajectories provides an implicit measure of the difference between predicted and observed outcomes. This difference serves as practical feedback to refine intervention strategies and improve the fit between model predictions and actual learning behavior. This result reinforces the importance of incorporating observed outcomes into subsequent decision cycles to progressively refine recommendation strategies. The system's observed behavior can be explained by several structural features of the proposed framework. First, the predictive model estimates performance based on correlations between behavioral traits and outcomes, rather than causal relationships. Therefore, improvements proposed through model-based recommendations may not fully translate into observed performance changes when underlying conditions differ. Second, the intervention mechanism assumes that recommended actions are performed consistently and with sufficient quality. In reality, differences in learner compliance levels, engagement intensity, and external learning conditions create uncertainty, which can lead to smaller observed changes than the model-based expected improvements. Third, current formulations treat the impact of behavioral changes as additive, whereas real-world learning processes are often non-linear and subject to diminishing returns, particularly for high-achieving learners. This can lead to a systematic overestimation of small improvements.

To improve the system, several approaches can be considered. The difficulty and feasibility of each recommendation type should be re-evaluated in light of real-world implementation to support more informed recommendation selection. Currently, the model is updated with new weekly learning data, but intervention logs and learner response data have not yet been incorporated systematically into the model. At this stage, these data are mainly used for post hoc monitoring and feedback analysis. Furthermore, to improve predictive quality, future work can experiment with more context-specific models and compare their performance across different learner groups, course modules, or instructional settings. For example, learners may be grouped by similar learning behavior patterns, prior performance levels, or course-specific activity profiles. Expanding the framework toward adaptive or personalized recommendation policies, where the usefulness and feasibility of actions are learned over time, may enhance the system's reliability across diverse learning scenarios.

5. Conclusions

This study presented an applied closed-loop learning analytics framework that integrates prediction, decision support, explainable AI, recommendation planning, and feedback logging into a unified workflow for process-oriented student assessment. By com-

binning weekly LMS-based behavioral indicators with comparative regression modeling and a SHAP-based explanation layer, the system can generate interpretable and instructor-reviewable recommendations at the learner–course–week level.

The results show that short-term learning performance can be predicted with reasonable stability in the current dataset. The comparative evaluation indicates that linear regression achieved the strongest predictive performance, while the previous-week score baseline also performed competitively, confirming the importance of recent performance history in week-ahead prediction. LightGBM did not achieve the best accuracy across all metrics, but it was retained for the explanation and recommendation-planning workflow because of its compatibility with SHAP-based feature attribution. These findings suggest that the value of the proposed system lies not in predictive superiority alone, but in connecting prediction, explanation, decision support, and feedback logging in an operational closed-loop process.

The intervention-planning analysis further illustrates how predicted scores and SHAP-based local explanations can be translated into feasible pedagogical recommendations. However, the observed differences between predicted and actual trajectories should be interpreted as preliminary feedback rather than causal evidence of intervention effectiveness. The system can support instructors by identifying risk-related behavioral factors and suggesting instructor-reviewable actions, but the actual effects of these recommendations depend on learner response, implementation quality, and instructional context.

The findings should also be interpreted in relation to the assessment context in which the system is deployed. In GenAI-aware courses or courses using more applied, authentic, project-based, oral, or AI-resistant assessment tasks, the meaning of LMS indicators and assessment-derived labels may change. Therefore, future deployment should document the assessment policy of each course and recalibrate the predictive features, decision thresholds, and recommendation mappings when the system is transferred to other disciplines, institutions, or assessment designs.

Future work will focus on strengthening empirical validation of the closed-loop process. Larger-scale studies across multiple courses, disciplines, and institutions are needed to examine the generalizability of the framework. Further research should also incorporate controlled or quasi-experimental designs to evaluate intervention effectiveness, refine recommendation costs and feasibility constraints using implementation data, and investigate ethical issues related to transparency, privacy, fairness, human oversight, and responsible use of AI in educational decision-making.

Author Contributions: Conceptualization, T.T.H.G. and Y.-J.R.; data curation, T.T.H.G.; formal analysis, T.T.H.G.; methodology, T.T.H.G. and Y.-J.R.; programming, T.T.H.G.; validation, Y.-J.R.; supervision, Y.-J.R.; review and editing, T.T.H.G. and Y.-J.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: Tay Nguyen University supported this research under grant number T2025–24CB. This work was supported by the Glocal University Project of Mokpo National University in 2026.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Viberg, O.; Hatakka, M.; Bälter, O.; Mavroudi, A. The current landscape of learning analytics in higher education. *Comput. Hum. Behav.* **2017**, *89*, 98–110. <https://doi.org/10.1016/j.chb.2018.07.027>.
2. Larrabee Sønderlund, A.; Hughes, E.; Smith, J. The efficacy of learning analytics interventions in higher education: A systematic review. *Br. J. Educ. Technol.* **2019**, *50*, 2594–2618. <https://doi.org/10.1111/bjjet.12720>.
3. Gašević, D.; Merceron, A. *The Handbook of Learning Analytics*; SoLAR: Beaumont, AB, Canada, 2022. <https://doi.org/10.18608/hla22>.
4. Chi, M.T.H.; Wylie, R. The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educ. Psychol.* **2014**, *49*, 219–243. <https://doi.org/10.1080/00461520.2014.965823>.
5. Ferguson, R.; Clow, D. Where is the evidence? A call to action for learning analytics. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*; ACM: New York, NY, USA, 2017; pp. 56–65. <https://doi.org/10.1145/3027385.3027396>.
6. Sailer, M.; Ninaus, M.; Huber, S.E.; Bauer, E.; Greiff, S. The End is the Beginning is the End: The closed-loop learning analytics framework. *Comput. Hum. Behav.* **2024**, *158*, 108305. <https://doi.org/10.1016/j.chb.2024.108305>.
7. Clow, D. The learning analytics cycle. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*; ACM: New York, NY, USA, 2012; pp. 134–138. <https://doi.org/10.1145/2330601.2330636>.
8. Wise, A.F.; Jung, Y. Teaching with Analytics : Towards a Situated Model of Instructional Decision-Making. *J. Learn. Anal.* **2019**, *6*, 53–69.
9. Krstić, L.; Aleksić, V.; Krstić, M. Artificial Intelligence in Education: A Review. In *Proceedings of the Proceedings TIE 2022*, Čačak, Serbia, 16–18 September 2022; pp. 223–228. <https://doi.org/10.46793/TIE22.223K>.
10. Akçapınar, G.; Altun, A.; Aşkar, P. Using learning analytics to develop early-warning system for at-risk students. *Int. J. Educ. Technol. High. Educ.* **2019**, *16*, 40. <https://doi.org/10.1186/s41239-019-0172-z>.
11. Katiyar, P.D.N.; Awasthi, V.K.; Pratap, R.; Mishra, K.; Shukla, N.; Singh, R.; Tiwari, M. Ai-Driven Personalized Learning Systems: Enhancing Educational Effectiveness. *Educ. Adm. Theory Pract.* **2024**, *30*, 11514–11524. <https://doi.org/10.53555/kuey.v30i5.4961>.
12. Tzimas, D.E.; Demetriadis, S.N. Impact of Learning Analytics Guidance on Student Self-Regulated Learning Skills, Performance, and Satisfaction: A Mixed Methods Study. *Educ. Sci.* **2024**, *14*, 92. <https://doi.org/10.3390/educsci14010092>.
13. Gupta, V.; Nyamapfene, A. Generative AI in Universities: Practices at UCL and Other Institutions, and the Path Forward. *Internet Ref. Serv. Q.* **2025**, *29*, 131–151. <https://doi.org/10.1080/10875301.2025.2453461>.
14. Gupta, V.; Nyamapfene, A. Integrating Generative AI in Academic Assessments: Practical Recommendations on AI Usage and Enhancing Knowledge Sharing. *Internet Ref. Serv. Q.* **2025**, *29*, 213–223. <https://doi.org/10.1080/10875301.2025.2456903>.
15. Cordero, J.; Torres-Zambrano, J.; Cordero-Castillo, A. Integration of Generative Artificial Intelligence in Higher Education: Best Practices. *Educ. Sci.* **2024**, *15*, 32. <https://doi.org/10.3390/educsci15010032>.
16. Ali, R.; Maroulis, J. Assessment validity in the age of generative artificial intelligence: A critical review. *Assess. Eval. High. Educ.* **2026**, 1–16. <https://doi.org/10.1080/02602938.2026.2683046>.
17. Siemens, G. Learning Analytics: The Emergence of a Discipline. *Am. Behav. Sci.* **2013**, *57*, 1380–1400. <https://doi.org/10.1177/0002764213498851>.
18. Lykourantzou, I.; Giannoukos, I.; Nikolopoulos, V.; Mpardis, G.; Loumos, V. Dropout prediction in e-learning courses through the combination of machine learning techniques. *Comput. Educ.* **2009**, *53*, 950–965. <https://doi.org/10.1016/j.compedu.2009.05.010>.
19. Gardner, J.; Brooks, C. Student success prediction in MOOCs. *User Model. User-adapt. Interact.* **2018**, *28*, 127–203. <https://doi.org/10.1007/s11257-018-9203-z>.
20. Maloney, S.; Axelsen, M.; Galligan, L.; Turner, J.; Redmond, P.; Brown, A.; Basson, M.; Lawrence, J. Using LMS Log Data to Explore Student Engagement with Coursework Videos. *Online Learn.* **2022**, *26*, 399–423. <https://doi.org/10.24059/olj.v26i4.2998>.
21. Boujmiraz, S.; Darhmaoui, H.; Drissi el Maliani, A. Predicting student performance: A comprehensive review of machine learning, deep learning, and explainable AI approaches. *Comput. Educ. Artif. Intell.* **2025**, *10*, 100548. <https://doi.org/10.1016/j.caeai.2026.100548>.
22. Zimmerman, B.J. Becoming a Self-Regulated Learner: An Overview. *Theory Into Pract.* **2002**, *41*, 64–70. https://doi.org/10.1207/s15430421tip4102_2.
23. Panadero, E. A review of self-regulated learning: Six models and four directions for research. *Front. Psychol.* **2017**, *8*, 422. <https://doi.org/10.3389/fpsyg.2017.00422>.
24. Dawson, S. A study of the relationship between student social networks and sense of community. *Educ. Technol. Soc.* **2008**, *11*, 224–238.
25. Wasserman, S.; Faust, K. *Social Network Analysis*; Cambridge University Press: Cambridge, UK, 1994. <https://doi.org/10.1017/CBO9780511815478>.

26. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. <https://doi.org/10.1214/aos/1013203451>.
27. Chen, T. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; ACM: New York, NY, USA, 2016.
28. Romero, C.; Ventura, S. Educational Data Mining and Learning Analytics : An Updated Survey. *WIREs Data Min. Knowl. Discov.* **2020**, *10*, e1355.
29. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements Of Statistical Learning*; Springer: New York, NY, USA, 2009; Volume 1. https://doi.org/10.1007/978-0-387-21606-5_7.
30. Breiman, L.E.O. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. <https://doi.org/10.1023/A:1010933404324>.
31. Zhang, L.; Wen, Z. A Transformer—Based Numerical Predictive Model of Student Academic Risk by Fusing Data from Multiple Sources. In *Proceedings of the 9th International Conference on Electronic Information Technology and Computer Engineering*; ACM: New York, NY, USA, 2025; pp. 1016–1021. <https://doi.org/10.1145/3766671.3766846>.
32. Adefemi, K.O.; Mutanga, M.B. A Robust Hybrid CNN–LSTM Model for Predicting Student Academic Performance. *Digital* **2025**, *5*, 16. <https://doi.org/10.3390/digital5020016>.
33. Airlangga, G. Predicting Student Performance Using Deep Learning Models: A Comparative Study of MLP, CNN, BiLSTM, and LSTM with Attention, MALCOM Indones. *J. Mach. Learn. Comput. Sci.* **2024**, *4*, 1561–1567. <https://doi.org/10.57152/malcom.v4i4.1668>.
34. Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **2017**, *2017*, 4766–4775.
35. Lundberg, S.M.; Erion, G.G.; Lee, S.-I. Consistent Individualized Feature Attribution for Tree Ensembles. *arXiv* **2019**, arXiv:1802.03888. <https://doi.org/10.48550/arXiv.1802.03888>
36. Karuru, P.; Suwarni, S.; Saddhono, K. Decision Support Systems in Education: A Revolutionary Approach to Teaching and Learning. *AL-ISHLAH J. Pendidik.* **2023**, *15*, 5015–5022. <https://doi.org/10.35445/alishlah.v15i4.4360>.
37. Alfredo, R.; Echeverria, V.; Jin, Y.; Yan, L.; Swiecki, Z.; Gašević, D.; Martinez-Maldonado, R. Human-centred learning analytics and AI in education: A systematic literature review. *Comput. Educ. Artif. Intell.* **2024**, *6*, 100215. <https://doi.org/10.1016/j.caeai.2024.100215>.
38. Vickers, A.J.; Elkin, E.B. Decision curve analysis: A novel method for evaluating prediction models. *Med. Decis. Mak.* **2006**, *26*, 565–574. <https://doi.org/10.1177/0272989X06295361>.
39. Srinivasan, V. AI & learning: A preferred future. *Comput. Educ. Artif. Intell.* **2021**, *3*, 100062. <https://doi.org/10.1016/j.caeai.2022.100062>.
40. Gunasekara, S.; Saarela, M. Systematic Literature Review on Explainable Learning Analytics and Educational Data Mining. *IEEE Access* **2025**, *13*, 214387–214408. <https://doi.org/10.1109/ACCESS.2025.3643645>.
41. Khosravi, H.; Shum, S.B.; Chen, G.; Conati, C.; Tsai, Y.-S.; Kay, J.; Knight, S.; Martinez-Maldonado, R.; Sadiq, S.; Gašević, D. Explainable Artificial Intelligence in education. *Comput. Educ. Artif. Intell.* **2022**, *3*, 100074. <https://doi.org/10.1016/j.caeai.2022.100074>.
42. ISO 8601-1:2019, Date and Time—Representations for Information Interchange—Part 1: Basic Rules. International Organization for Standardization: Geneva, Switzerland, 2019.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.