OPEN ACCESS
International Journal of Innovations in Science & Technology
RESEARCH & INNOVATION DIVISION
IJIST

Human–Computer Interaction Patterns in E-Learning: Insights from Learning Behavior Mining for Performance Prediction

Fasih Iqbal[1], Hassam Iqbal[2]
[1]Department of Artificial Intelligence, Capital University of Science and Technology, Islamabad, Pakistan
[2]Faculty of Medicine, Dentistry and Health Sciences, the University of Melbourne, Victoria, Australia
*Correspondence: fasihmughal123@gmail.com, hassaamughal@gmail.com

Citation | Iqbal. F, Iqbal. H, "Human–Computer Interaction Patterns in E-Learning: Insights from Learning Behavior Mining for Performance Prediction", IJIST, Special Issue pp 757-776, May 2026

Received| April 11, 2026 **Revised|** May 17, 2026 **Accepted|** May 19, 2026 **Published|** May 21, 2026.

Human–computer interaction (HCI) plays a significant role in shaping learner engagement and outcomes in e-learning environments. Existing learning analytics approaches often rely on aggregated behavioral features, thereby limiting interpretability and weakening the connection between prediction and learner interaction processes. This study proposes an HCI-informed learning behavior representation that captures navigation dynamics, temporal interaction characteristics, engagement signals, and self-regulation behaviors from e-learning interaction logs. Using the Open University Learning Analytics Dataset (OULAD), logistic regression, random forest, and gradient boosting models were evaluated under static and HCI-informed feature settings. The proposed HCI-informed representation consistently improved predictive performance across all evaluated models. The best-performing configuration, Random Forest with HCI-informed features, achieved an accuracy of 0.858, an F1-score of 0.872, and an AUC of 0.925, and AUC of 0.925, compared with 0.796, 0.810, and 0.878 respectively for the static baseline. Statistical validation further demonstrated significant AUC improvements across all models ($p < 0.001$). Beyond prediction, the proposed representation provides interpretable evidence linking navigation regularity, temporal engagement, and self-regulation patterns with learner outcomes. These findings demonstrate that learning analytics can become more explanatory and actionable when interaction behavior is modeled from an HCI-informed perspective.

Keywords: Human–Computer Interaction, Learning Analytics, Educational Data Mining, E-Learning, Performance Prediction

Introduction:

The rapid growth of e-learning platforms has transformed how learners interact with educational content, assessments, and peers. Modern learning management systems (LMSs) continuously record detailed logs of learner interactions, including navigation actions, content access, and assessment activities. These interaction traces provide valuable opportunities to understand learning behaviors and predict academic performance at scale. Consequently, learning analytics and educational data mining have become central to the development of data-driven educational technologies [1][2].

Prior research has demonstrated that learner interaction data can be used to predict academic outcomes such as course completion, grades, and dropout risk. Feature-based learning behavior mining approaches have shown promising results by extracting aggregated measures of learner activity and applying classical machine learning models. However, many existing approaches primarily prioritize predictive accuracy while paying limited attention to the human–computer interaction processes that generate these data. As a result, the relationship between predicted outcomes and underlying learner behaviors often remains opaque, reducing the interpretability and practical utility of such systems.

Recent studies have increasingly emphasized explainability, intervention support, and human-centered learning analytics in student performance prediction research [3][4][5][6]. Contemporary machine learning approaches have demonstrated strong predictive capability using online learning interaction data, but recent reviews highlight that many systems still lack interpretability and actionable behavioral insight [7]. Human-centered learning analytics research has further emphasized the importance of transparency, stakeholder trust, fairness, and educational usability in predictive systems [8][9][10]. These developments indicate a growing shift from prediction-focused analytics toward explainable and adaptive learning analytics frameworks that support meaningful educational decision-making [11].

From a human–computer interaction (HCI) perspective, learning is inherently a dynamic and interactive process. Learners engage in sequences of actions such as navigating between resources, revisiting content, attempting assessments, and regulating their study behavior over time. Aggregated behavioral features may obscure these interaction dynamics and fail to capture important temporal and structural aspects of learning behavior. Understanding how specific interaction patterns reflect engagement, confusion, persistence, or self-regulation is essential for designing adaptive and human-centered e-learning systems.

This paper aims to bridge the gap between learning behavior mining and human–computer interaction by investigating how fine-grained interaction patterns can be systematically represented and interpreted for learning performance prediction. Using a publicly available e-learning dataset, we propose an HCI-informed learning behavior representation that organizes learner interactions into interpretable behavioral constructs, including navigation behavior, temporal interaction characteristics, engagement signals, and self-regulation patterns. We empirically evaluate the predictive value of these features and compare them with traditional static feature representations commonly used in prior work.

Research Objectives:

The present study is guided by the following research objectives:

To develop an HCI-informed representation of learner interaction behavior using clickstream and temporal interaction data.

To compare the predictive performance of traditional static feature representations and HCI-informed behavioral features for learner performance prediction.

To investigate whether interaction-aware features provide interpretable insights related to navigation behavior, temporal engagement, and self-regulation.

To explore the implications of HCI-informed learning analytics for adaptive and human-centered e-learning system design.

Novelty and Research Contribution:

The novelty of this study lies in reframing learning performance prediction as a human–computer interaction-informed behavior interpretation problem rather than solely a classification task. Unlike conventional learning analytics approaches that rely primarily on aggregated activity counts, the proposed framework organizes learner interaction traces into interpretable HCI constructs, including navigation behavior, temporal engagement, interaction diversity, and self-regulation signals.

The study makes three important contributions. First, it introduces an HCI-informed feature representation that captures both the structural and temporal dimensions of learner interaction behavior. Second, it demonstrates that interaction-aware representations significantly improve predictive performance across multiple machine learning models. Third, the framework enhances interpretability by linking predictive outcomes to understandable learner behaviors, thereby supporting more transparent and actionable learning analytics systems.

We propose an HCI-informed framework for learning behavior mining that systematically captures fine-grained learner interaction patterns.

We demonstrate that interaction-aware and temporal behavior features improve learning performance prediction compared with traditional static feature-based approaches.

We provide interpretable insights into how specific human–computer interaction patterns relate to effective and ineffective learning, offering actionable implications for the design of adaptive e-learning systems.

The remainder of this paper is organized as follows. Section II reviews related work in learning analytics and human–computer interaction. Section III describes the dataset and interaction logging process. Section IV presents the proposed HCI-informed learning behavior representation. Section V details the experimental setup and evaluation methodology. Section VI reports the experimental results, followed by a discussion of HCI insights and design implications in Section VII. Finally, Section VIII outlines limitations and future research directions, and Section IX concludes the paper.

Related Work:

Learning Analytics and Performance Prediction:

Learning analytics has attracted significant attention due to the increasing availability of fine-grained interaction data from learning management systems. Early studies demonstrated that learner activity logs, such as resource access frequency and assessment attempts, can be used to predict academic performance and dropout risk. Traditional approaches primarily relied on aggregated behavioral features and classical machine learning models, including logistic regression, decision trees, and random forests [12][13][14][15][16].

More recent work has explored richer feature representations to capture learning behavior more effectively. Feature-based learning behavior mining approaches extract statistics related to content access, assessment engagement, and temporal activity patterns to improve prediction accuracy. Among these, effective behavior characterization models have demonstrated that carefully engineered behavioral features can yield strong predictive performance on e-learning datasets [17]. However, such approaches largely emphasize predictive performance while providing limited insight into the interaction processes that generate these behavioral patterns.

Recent learning analytics research has increasingly focused on explainability, adaptive educational support, and human-centered predictive systems [18][9][19]. Contemporary educational data mining studies have shown that interpretable behavioral representations can improve educational usability and intervention quality while maintaining strong predictive capability [20][21]. In parallel, self-regulated learning analytics research has highlighted the importance of temporal engagement consistency, persistence, and behavioral organization as indicators of learner success [22][23].

Learning Behavior Mining in E-Learning Environments:

Learning behavior mining aims to extract meaningful patterns from learner interaction data to understand how students engage with digital learning environments. Prior studies have analyzed clickstream data, time-on-task, and assessment attempts to identify behavioral patterns associated with successful learning outcomes [24]. Temporal aspects of learning behavior have also been shown to be important, particularly for early performance prediction and dropout detection.

Sequential modeling approaches, including hidden Markov models and recurrent neural networks, have been applied to capture learning trajectories over time. While these methods can model complex temporal dependencies, they often prioritize predictive performance over interpretability, making it difficult to translate model outputs into actionable educational insights.

Human–Computer Interaction Perspectives:

From a human–computer interaction perspective, learning is fundamentally shaped by how users interact with digital interfaces. Interaction design, navigation complexity, and feedback mechanisms play important roles in influencing learner engagement and cognitive processes [25]. Studies in this area emphasize that interaction traces reflect not only activity levels but also underlying cognitive and behavioral states.

Despite these insights, HCI perspectives remain underutilized in many learning analytics models. Interaction data are often treated as abstract numerical inputs without explicit consideration of their human-centered interpretation. Integrating HCI principles into learning behavior mining is therefore essential for improving both the interpretability and practical utility of predictive models [26].

Explainability in Educational Data Mining:

Interpretability has become an increasingly important concern in educational data mining, particularly for systems intended to support instructors and learners [7][27]. Black-box models may achieve high predictive accuracy but often fail to provide explanations that are meaningful to human stakeholders. Recent work has emphasized the importance of explainable artificial intelligence approaches in educational settings [7]. Model-agnostic explanation techniques and interpretable feature representations have been proposed to improve transparency and trust in predictive systems [27]. However, explainability is frequently treated as a post-hoc addition rather than a core design principle. This limitation motivates the development of learning behavior representations that are inherently interpretable and grounded in human–computer interaction concepts.

Research Gap:

In summary, existing learning performance prediction approaches demonstrate strong predictive capabilities but often rely on static feature representations and offer limited interpretability from a human–computer interaction perspective. There remains a lack of systematic frameworks that explicitly connect fine-grained interaction patterns with learning behavior mining and performance prediction across temporal scales [28]. This study addresses this gap by proposing an HCI-informed learning behavior representation that captures interaction dynamics while providing interpretable insights into learner behavior.

Dataset and interaction logging:

Dataset Description:

This study uses the Open University Learning Analytics Dataset (OULAD), a publicly available dataset that contains detailed records of learner interactions within a virtual learning environment [28]. OULAD has been widely used in learning analytics research and provides a comprehensive view of learner behavior across multiple courses and presentations. The dataset includes anonymized information on learner demographics, interaction logs, and assessment outcomes, enabling reproducible and ethically compliant research. Each course in OULAD is offered in multiple presentations, corresponding to specific academic terms. Learner interactions are recorded throughout the duration of each presentation, capturing fine-grained activity traces such as resource access events and assessment participation. These characteristics make OULAD well-suited for analyzing human–computer interaction patterns and their relationship with learning performance. Table 1 summarizes recent learning analytics and student performance prediction approaches and highlights the relevance of the proposed HCI-informed framework.

Table 1. Comparison of recent learning analytics and student performance prediction approaches.

Study	**Main Focus**	**Feature Type**	**Limitation**	**Relevance to This Study**
Albreiki et al. (2021)	Machine learning for student performance prediction	Academic and online interaction features	LimitedHCI interpretation	Motivates interpretable interaction-aware learning analytics
Dimitriadis et al. (2021)	Human-centered actionable learning analytics	Human-centered design principles	Limited predictive feature modeling	Supports HCI-informed interpretation and explainability
Recent explainable learning analytics studies	Explainable educational AI systems	Transparent prediction frameworks	Limited temporal interaction modeling	Supports interpretable behavior-based prediction
Recent adaptive learning analytics research	Adaptive intervention systems	Engagement and intervention signals	Limited behavioral structure analysis	Supports adaptive HCIinformed learning systems
This study	HCI-informed learning behavior mining	Navigation, temporal, engagement, and self-regulation features	Single dataset evaluation	Bridges predictive modeling and interpretable HCI behavior analysis

Although OULAD provides a rich and well-established benchmark for learning analytics research, the use of a single dataset imposes natural limits on external validity. Interaction patterns observed in one virtual learning environment may not fully reflect learner behavior in platforms with different interface designs, pedagogical structures, or learner populations. For this reason, the present study should be interpreted as an initial empirical validation of the proposed HCI-informed representation rather than a claim of universal generalizability.

Interaction Logging and Virtual Learning Environment:

Learner interactions in OULAD are captured through the virtual learning environment (VLE) and recorded as timestamped events. Each interaction corresponds to a learner accessing a specific learning resource, such as lecture materials, quizzes, forums, or supplementary content. The dataset records the type of resource accessed, the time of interaction, and the frequency of access, providing a detailed representation of learner engagement with the interface.

The interaction logs enable the reconstruction of learner navigation paths and temporal activity patterns. By analyzing sequences of interactions over time, it is possible to capture how learners move through the learning environment, revisit materials, and allocate effort across different learning activities. These interaction traces serve as the foundation for mining human–computer interaction behaviors and understanding learning dynamics beyond simple activity counts.

It is important to note, however, that interaction logs capture only observable learner actions within the digital environment. They do not directly represent offline study behavior, attention, motivation, or deeper cognitive engagement. Consequently, the interaction traces used in this study should be understood as informative behavioral proxies rather than complete representations of the learning process.

Assessment and Performance Measures:

Learner performance in OULAD is measured through a combination of formative and summative assessments, including quizzes, assignments, and final examinations. Each assessment is associated with a submission date, weight, and score. In this study, the final course outcome is used as the primary target variable for performance prediction, following common practice in prior learning analytics research. The outcome is derived from assessment scores and represents overall learner achievement at the end of the course.

In addition to final outcomes, intermediate assessment information is used to support early-stage performance prediction. This allows the analysis of how interaction behaviors observed during the early phases of a course relate to eventual learning outcomes, which is critical for timely intervention and support.

Data Preprocessing and Ethical Considerations:

All data used in this study are fully anonymized and publicly available. No personally identifiable information is included in the dataset. To ensure data quality and consistency, preprocessing steps are applied prior to feature extraction. These steps include filtering incomplete learner records, aligning interaction events to a common temporal scale, and normalizing interaction counts across different resource types.

Ethical considerations are addressed by adhering to the original dataset usage guidelines and ensuring that all analyses focus on aggregated behavioral patterns rather than individual identification. The use of a publicly available dataset enables transparency, reproducibility, and comparability with prior work in learning analytics and educational data mining.

Formal Dataset Representation:

Let $U = \{u_1, u_2, \ldots, u_N\}$ denote the set of learners and let $E_i = \{e_{i1}, e_{i2}, \ldots, e_{iT_i}\}$ represent the sequence of interaction events generated by learner u_i during a course. Each interaction event e_{ij} is defined as a tuple:

$$eij = (rij,\ tij) \quad (1)$$

where r_{ij} denotes the accessed learning resource and t_{ij} denotes the timestamp of the interaction.

These interaction sequences form the basis for extracting human–computer interaction features that summarize learner behavior over time.

HCI-Informed Learning Behavior Representation:

Overview of the HCI-Informed Learning Behavior Representation:

To capture meaningful human–computer interaction patterns in e-learning environments, we propose an HCI-informed learning behavior representation that explicitly models how learners interact with the virtual learning environment over time. Rather than relying solely on aggregated activity counts, the proposed approach organizes interaction data into interpretable behavioral constructs grounded in principles from human–computer interaction and learning sciences.

The proposed representation is designed to balance predictive capability and interpretability. It captures both the intensity and structural organization of learner interactions, enabling analysis of navigation dynamics, temporal engagement patterns, and self-regulation behaviors. This structured representation provides a foundation for understanding how interaction behaviors relate to learning outcomes while supporting the development of human-centered predictive models.

Problem Formulation:

Let i denote a learner and let $\mathbf{x}_i \in \mathrm{R}^d$ represent the HCI-informed feature vector extracted from the learner's interaction logs, where d denotes the number of interaction-aware behavioral features. These features summarize how the learner interacts with the virtual learning environment in terms of navigation behavior, temporal engagement, and self-regulation.

Each learner is associated with a performance outcome label y_i, representing the final course outcome. The learning performance prediction task is formulated as a supervised classification problem in which a predictive model $f(\cdot)$ maps interaction features to an outcome prediction:

$$\hat{y}_i = f(\mathbf{x}_i) \quad (2)$$

The objective is to learn a function f that accurately predicts learner performance based on human–computer interaction patterns observed during the learning process.

Behavioral Feature Taxonomy:

Learner interaction data are organized into four primary behavioral categories, each reflecting a distinct aspect of human–computer interaction within the learning environment.

Feature Aggregation Across Interaction Events:

For each learner u_i, the extracted interaction events are aggregated into an HCI-informed feature vector $\mathbf{x}_i$. Each feature is computed as a function over the learner's interaction sequence:

$$x_{ik} = g_k(\mathrm{E}_i) \quad (3)$$

Where $g_k(\cdot)$ denotes a feature extraction function corresponding to a specific interaction behavior, such as navigation frequency, temporal regularity, or engagement consistency.

This aggregation process enables the transformation of variable-length interaction sequences into fixed-length and interpretable feature representations suitable for predictive modeling. Figure 1 illustrates the temporal aggregation process used to transform learner interaction events into fixed-length HCI-informed feature representations.

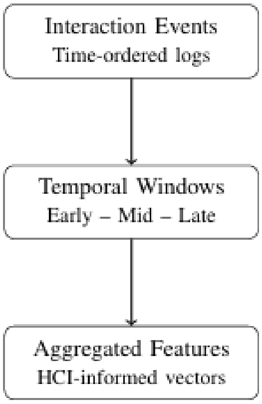

Figure 1. Temporal aggregation of learner interaction events into fixed-length HCI-informed feature representations.

Navigation Behavior:

Navigation behavior captures how learners move through the virtual learning environment. Features in this category describe the diversity, frequency, and structure of resource access, reflecting how learners explore and revisit learning materials. Examples include the number of unique resources accessed, revisit frequency, and navigation entropy, which quantifies the variability of navigation paths. High navigation entropy may indicate exploratory behavior or potential confusion, whereas focused navigation patterns may reflect goal-directed learning.

Temporal Interaction Characteristics:

Temporal interaction features model the timing and rhythm of learner activity. These features capture inter-event timing, session duration, and activity burstiness, providing insight into how learners distribute effort over time. Regular study rhythms may reflect effective self-regulation, whereas highly irregular or sporadic activity patterns may indicate disengagement or poor time management behavior.

Engagement Signals:

Engagement-related features characterize the depth and intensity of learner interaction with course content. These include measures such as total interaction frequency, resource access diversity, and consistency of engagement across learning sessions. Sustained and consistent engagement is commonly associated with positive learning outcomes, whereas shallow or declining engagement may indicate increased risk of poor performance.

Self-Regulation and Persistence:

Self-regulation features capture learners' ability to plan, monitor, and adapt their learning behavior. These include indicators such as early versus late engagement with assessments, repeated attempts on learning activities, and consistency of participation across the course timeline. Persistent interaction patterns and early engagement with course materials often reflect proactive learning strategies and learner resilience.

The proposed feature taxonomy is intended to provide an interpretable and practically useful abstraction of learner behavior rather than an exhaustive representation of the learning process. Real learning behavior is inherently multidimensional and context-dependent, and not all aspects of that complexity can be captured through a finite set of behavioral descriptors. Nevertheless, the structured feature grouping adopted in this study provides a transparent framework for connecting interaction traces with meaningful HCI constructs.

HCI-Informed Learning Behavior Mining Pipeline:

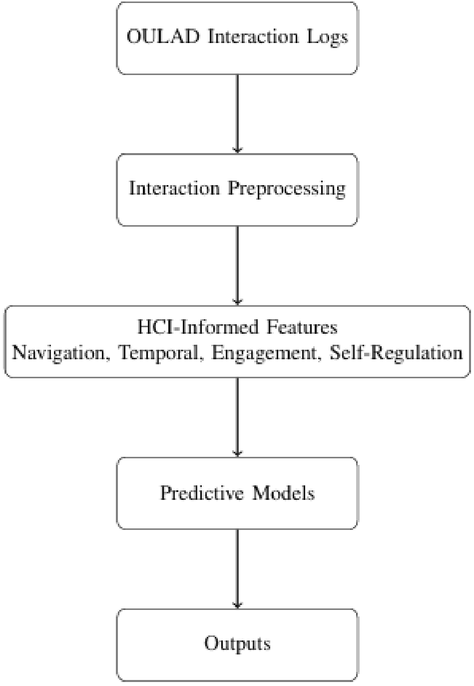

Figure 2. Overview of the HCI-informed learning behavior mining pipeline.

Figure 2 illustrates the overall workflow of the proposed framework. Raw learner interaction logs are first preprocessed to remove incomplete records and organize events chronologically. The system then extracts HCI-informed behavioral features related to navigation behavior, temporal engagement, interaction diversity, and self-regulation. These features are subsequently used to train predictive machine learning models for learner performance prediction and early risk identification. The figure emphasizes the transformation of raw interaction traces into interpretable behavioral representations that support both predictive performance and educational insight generation.

Algorithm 1. HCI-informed feature extraction from learner interaction logs

```
Require: Learner interaction logs E
Ensure: HCI-informed feature matrix X
for each learner ui do
    Sort interaction events chronologically
    Remove incomplete or invalid records
    Compute navigation features:
        resource revisits, navigation entropy
    Compute temporal interaction features:
        session regularity, inter-event timing
    Compute engagement features:
        activity diversity, interaction frequency
    Compute self-regulation features:
        early engagement, persistence indicators
    Aggregate all extracted features into feature vector xi
end for
Return feature matrix X
```

Temporal Aggregation Strategy:

To incorporate temporal dynamics while maintaining interpretability, learner interaction data are aggregated within predefined temporal windows corresponding to early, middle, and late stages of the course. This strategy enables analysis of how interaction behaviors evolve over time and supports early-stage performance prediction. By comparing behavioral representations across temporal windows, it becomes possible to identify changes in engagement and navigation patterns that precede successful or unsuccessful learning outcomes.

Feature Normalization and Representation:

All behavioral features are normalized to account for differences in course length, resource availability, and learner activity levels. Normalization ensures that features remain comparable across learners while reducing bias introduced by course-specific characteristics. The final learning behavior representation is constructed as a structured feature vector that preserves the semantic grouping of interaction behaviors, facilitating both predictive modeling and interpretability.

Because the proposed representation relies on fine-grained interaction logs, the computational cost of feature extraction may increase as the number of learners, events, and temporal windows grows. In the present study, the emphasis is placed on interpretability and analytical clarity rather than large-scale deployment; however, scalability remains an important consideration for operational implementation.

Modeling Considerations:

The proposed HCI-informed learning behavior representation is designed to be model-agnostic. It can be used with a range of machine learning models, including logistic regression, tree-based methods, and ensemble models. This flexibility enables systematic comparison between traditional static feature-based approaches and interaction-aware representations, facilitating empirical evaluation of the benefits of incorporating HCI principles into learning performance prediction.

By grounding feature design in human–computer interaction concepts, the proposed methodology provides a transparent and interpretable framework for learning behavior mining. This approach supports both accurate performance prediction and meaningful insight into learner interaction processes, addressing key limitations of prior feature-based learning analytics models.

Experimental Setup:

Experimental Design:

The experimental setup is designed to evaluate the effectiveness of the proposed HCI-informed learning behavior representation for predicting learner performance. We compare interaction-aware and temporal behavior features against traditional static feature representations commonly used in prior learning analytics studies. All experiments are conducted using the same dataset splits, evaluation metrics, and modeling procedures to ensure fair comparison.

Prediction Task:

The primary prediction task is learner performance prediction at the end of the course. Following common practice in learning analytics research, learner performance is formulated as a supervised classification problem, where the goal is to predict final course outcome based on interaction behaviors observed during the course. To support early intervention analysis, additional experiments evaluate prediction performance using interaction data from early stages of the course only.

Baseline Feature Representation:

Baseline models use static, aggregated behavioral features that summarize learner activity across the entire observation period. These features include total interaction counts, total number of accessed resources, and cumulative assessment participation. This representation reflects commonly used feature engineering strategies in existing learning behavior mining approaches and serves as a reference point for evaluating the proposed HCI-informed representation.

HCI-Informed Feature Representation:

The proposed feature representation incorporates navigation behavior, temporal interaction characteristics, engagement signals, and self-regulation indicators as described in Section IV. Features are computed within predefined temporal windows to capture changes in interaction patterns over time. This representation preserves the semantic structure of learner interactions and enables analysis of both interaction intensity and behavioral dynamics.

Predictive Models:

To ensure interpretability and reproducibility, we evaluate a set of widely used machine learning models. These include logistic regression as a linear baseline and tree-based ensemble methods such as random forests and gradient boosting machines. These models provide a balance between predictive performance and interpretability and are commonly used in educational data mining research.

All models are trained using the same feature sets and optimized using standard hyperparameter tuning procedures. Model selection is performed based on comparative predictive performance across validation experiments while maintaining consistency in evaluation settings.

At the same time, early-stage prediction presents an inherent methodological challenge because learner interaction traces are relatively sparse during the initial phase of a course. The evaluation therefore considers early prediction as a practically important but more uncertain setting, where robustness and interpretability are especially valuable. Table 2 summarizes the model configurations and hyperparameter settings used in the experimental evaluation.

All experiments were conducted using a stratified 80/20 train–test split to preserve class balance between successful and at-risk learners. Hyperparameter settings were selected based on preliminary validation experiments and commonly used configurations in educational data mining literature. Performance was evaluated using Accuracy, F1-score, and Area under the Receiver Operating Characteristic Curve (AUC).

Table 2. Model configuration and hyperparameter settings.

Model	**Hyperparameter settings**
Logistic Regression	Penalty = L2, max_iter = 1000, class_weight = balanced
Random Forest	n_estimators = 300, max_depth = 10, class_weight = balanced, random_state = 42
Gradient Boosting	n_estimators = 200, learning_rate = 0.05, max_depth = 3, random_state = 42

Results:

Overall Prediction Performance:

This section presents experimental results evaluating the effectiveness of the proposed HCI-informed learning behavior representation for learner performance prediction. We compare the predictive performance of models trained using traditional static feature representations with those using the proposed interaction-aware and temporal behavioral features. Table 3 reports the predictive performance of static and HCI-informed feature representations across all evaluated models.

Table 3. Performance comparison between static and HCI-informed feature representations.

Model	Feature Set	Accuracy	F1-score	AUC
Logistic Regression	Static	0.779	0.779	0.867
Logistic Regression	HCI-informed	0.836	0.842	0.914
Random Forest	Static	0.796	0.810	0.878
Random Forest	HCI-informed	0.858	0.872	0.925
Gradient Boosting	Static	0.794	0.812	0.874
Gradient Boosting	HCI-informed	0.855	0.869	0.922

Table 4 presents the statistical validation results for AUC improvements obtained using HCI-informed features.

Table 4. Statistical validation of AUC improvements using HCI-informed features

Model	**Mean AUC Gain**	**95% CI Lower**	**95% CI Upper**	**p-value**
Logistic Regression	0.0437	0.0421	0.0454	< 0.001
Random Forest	0.0451	0.0437	0.0466	< 0.001
Gradient Boosting	0.0451	0.0427	0.0475	< 0.001

The statistical validation results indicate that the observed improvements obtained using HCI-informed features are unlikely to be due to random variation. Across all evaluated models, the mean AUC improvement remained consistently positive, with narrow confidence intervals and statistically significant paired comparisons ($p < 0.001$). These findings support the reliability and robustness of the proposed interaction-aware feature representation.

Figure 3 visually demonstrates the consistent performance gains achieved using HCI-informed behavioral features across all evaluated machine learning models. The largest improvements are observed for ensemble-based approaches, particularly Random Forest and Gradient Boosting, suggesting that interaction-aware features capture non-linear behavioral relationships that are not adequately represented through static activity aggregation alone.

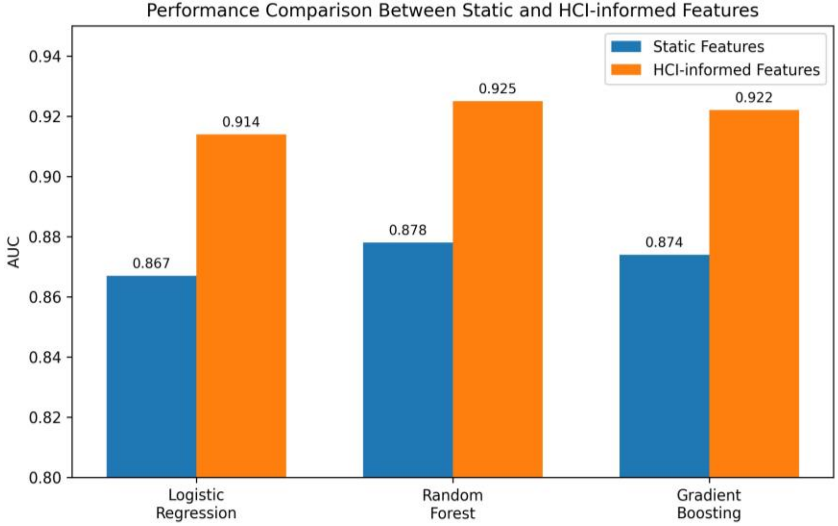

Figure 3. Comparison of AUC performance between static feature representations and the proposed HCI-informed feature representation across all evaluated machine learning models.

Across all evaluated models, the HCI-informed feature representation consistently outperformed the static baseline in terms of predictive accuracy and discriminative ability. Improvements were observed across multiple evaluation metrics, including accuracy, F1-score, and area under the receiver operating characteristic curve (AUC), indicating that incorporating fine-grained interaction patterns provides additional predictive signal beyond aggregated activity counts.

These improvements suggest that interaction-aware features capture aspects of learner behavior that are not represented in conventional activity-based metrics. While static features summarize the overall volume of learner activity, the proposed representation incorporates structural properties of interaction, including navigation patterns and temporal regularity. As a result, models trained on these features are better able to differentiate learners who interact strategically with course materials from those who engage in fragmented or irregular ways.

The observed improvement is meaningful because it suggests that performance prediction benefits not merely from a larger number of features, but from features that preserve the structure of learner interaction. In other words, the improvement appears to arise from modeling how learners engage with the platform rather than simply how often they interact with it. This distinction is important because it shifts the analytical focus from activity volume toward behavioral organization.

Comparison with Static Feature-Based Models:

Models trained on static aggregated features achieved reasonable baseline performance, consistent with findings reported in prior learning analytics studies. However, these models showed limited sensitivity to variations in learner interaction behavior over time. In contrast, models leveraging the proposed HCI-informed representation demonstrated improved performance, particularly in distinguishing learners with similar overall activity levels but different interaction strategies.

Tree-based ensemble models benefited most strongly from the proposed feature representation, suggesting that interaction-aware features capture non-linear relationships between learner behavior and performance outcomes. These findings highlight the limitations of purely static feature aggregation and underscore the importance of modeling interaction dynamics explicitly.

This pattern is consistent with the ability of ensemble models to capture complex relationships between behavioral features. Interaction-aware representations introduce dependencies between variables such as navigation structure, temporal engagement consistency, and study rhythm. Tree-based methods are better suited to exploit these relationships, which may explain their stronger performance relative to simpler linear models under the proposed representation.

This comparison further clarifies the practical value of the proposed approach. Static feature summaries may treat two learners as behaviorally similar if they generate comparable numbers of clicks or access similar numbers of resources, even when their interaction strategies differ substantially. By incorporating navigation structure and temporal regularity, the HCI-informed representation is better able to distinguish purposeful engagement from fragmented or inconsistent interaction behavior.

The observed improvements suggest that learner success is influenced not only by the quantity of interaction activity, but also by the organization and consistency of interaction behavior. HCI-informed features capture structural aspects of engagement that are largely overlooked by static aggregation methods. In particular, navigation entropy and temporal regularity appear to provide meaningful signals regarding learner persistence, study organization, and self-regulation.

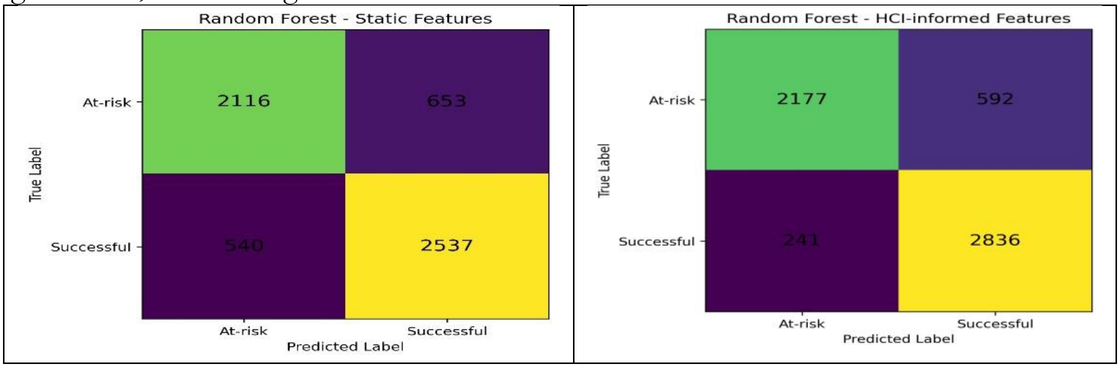

Figure 4. Confusion matrix comparison for Random Forest using static features (left) and HCI-informed features (right). The HCI-informed representation demonstrates improved classification performance for both at-risk and successful learners.

Figure 4 illustrates the classification behavior of the Random Forest model under static and HCI-informed feature settings. The HCI-informed representation produces more accurate classification across both learner categories, particularly reducing misclassification of at-risk learners. This improvement suggests that interaction-aware behavioral features provide stronger discriminatory information than aggregate activity measures alone.

Early-Stage Performance Prediction:

To evaluate the ability of the proposed approach to support early intervention, additional experiments restricted input features to interaction data collected during the early phase of the course. Under these conditions, models using HCI-informed features maintained substantially higher predictive performance than baseline models.

This finding should, however, be interpreted alongside the known difficulty of early prediction. During the initial phase of a course, the number of observed interactions is typically limited, increasing uncertainty and reducing predictive stability. The fact that HCI-informed features remain informative under these conditions suggests that behavioral structure, rather than activity volume alone, contributes to the early identification of learners who may require support.

The results indicate that early interaction patterns, including navigation behavior, temporal engagement rhythms, and initial self-regulation signals, provide meaningful indicators of eventual learning outcomes. This finding is particularly important for practical deployment because it enables timely identification of learners who may benefit from targeted support or intervention.

Feature Importance and Interpretability Analysis:

An analysis of feature importance revealed that interaction-aware features contributed more strongly to prediction outcomes than simple activity counts. Navigation-related features, including resource revisit frequency and navigation entropy, emerged as key predictors, suggesting that how learners move through the learning environment is closely related to learning success.

Temporal interaction characteristics, such as regularity of study sessions and inter-event timing, also played prominent roles, reflecting the importance of consistent engagement and self-regulated learning behavior. These findings support the interpretability of the proposed feature representation and demonstrate how human–computer interaction patterns can be meaningfully linked to learning performance.

Taken together, the results suggest that effective learners tend to exhibit structured interaction behavior rather than merely higher activity levels. Consistent engagement rhythms and purposeful navigation patterns indicate deliberate study strategies, whereas highly irregular interaction sequences may reflect uncertainty, confusion, or disengagement. This behavioral distinction highlights the advantage of modeling interaction dynamics instead of relying solely on aggregate activity measures.

From a qualitative perspective, these findings are equally important. The strongest predictors are not arbitrary machine learning signals; rather, they correspond to behaviors that can be interpreted by instructors and learning designers. For example, repeated revisits to key resources may reflect sustained effort or difficulty resolution, whereas irregular session timing may indicate unstable study habits. This interpretability increases the practical usefulness of the model because it connects predictive outcomes to understandable learner behaviors.

Summary of Results:

Overall, the experimental results demonstrate that incorporating HCI-informed and temporal learning behavior features leads to substantially improved predictive performance compared with traditional static feature representations. The strongest performance was achieved using Random Forest with HCI-informed features, reaching an accuracy of 0.858, F1-score of 0.872, and AUC of 0.925. Statistical validation further confirmed that the observed improvements were significant across all evaluated models. Beyond predictive gains, the proposed framework provides interpretable insights into learner interaction processes, supporting its use in both analytical and human-centered educational applications.

HCI Insights and Design Implications:

Interpreting Learning Behavior through an HCI Lens:

The results of this study demonstrate that learner interaction data contain rich signals that extend beyond simple measures of activity volume. From a human–computer interaction perspective, these signals reflect how learners perceive, navigate, and regulate their interaction with the learning environment. By organizing interaction features into interpretable behavioral constructs, the proposed approach enables a more nuanced understanding of learning processes that are often obscured by static feature aggregation.

Navigation-related behaviors emerge as particularly informative. High navigation entropy and frequent revisits to learning resources may indicate exploratory learning strategies, but they may also signal confusion or difficulty locating relevant content. Conversely, more structured navigation patterns suggest goal-oriented interaction and effective use of the interface. These findings highlight the importance of considering not only how much learners interact with the system, but also how they traverse and engage with its structure.

These observations reinforce the idea that interaction structure carries meaningful information about learner intent and strategy. Two learners may generate similar numbers of interactions while exhibiting very different navigation patterns. One learner may follow a coherent progression through course materials, whereas another may repeatedly jump between resources without clear direction. By capturing these structural differences, the proposed representation provides a richer perspective on learner engagement than activity counts alone.

Temporal interaction characteristics further reveal differences in learner engagement and self-regulation. Regular study rhythms and consistent session timing are associated with positive learning outcomes, whereas highly irregular or bursty interaction patterns may reflect procrastination or surface-level engagement. Such temporal patterns provide insight into learners' time-management strategies and their ability to sustain attention throughout the course duration.

This perspective shifts the interpretation of learner data away from a purely predictive paradigm toward a more explanatory framework. Instead of asking only whether a learner is likely to succeed, the proposed framework helps explain how different styles of interaction may support or hinder learning. This explanatory dimension is essential if predictive systems are to inform real educational decisions rather than function as isolated black-box classifiers.

Practical Implications for Human-Centered E-Learning Design:

The findings of this study have direct implications for the design of human-centered e-learning systems. First, interaction-aware analytics can support adaptive interface design by identifying navigation behaviors that signal confusion or disengagement. Learning platforms could use these signals to dynamically recommend relevant resources, simplify navigation pathways, or provide contextual guidance when learners exhibit inefficient navigation behavior.

Second, temporal engagement signals can inform the design of interventions that promote self-regulated learning. For example, systems may detect irregular study patterns early in a course and provide personalized prompts, planning tools, or reminders to encourage more consistent engagement. Such interventions align with HCI principles that emphasize timely feedback and learner support without imposing excessive cognitive load.

Importantly, these behavioral signals do not require intrusive monitoring or complex sensing technologies. They can be derived directly from interaction logs already collected by most learning management systems. This makes the proposed framework practical for real-world educational deployment, where scalable and minimally invasive analytics solutions are generally preferred.

Third, interpretable learning behavior representations enhance transparency and trust in predictive systems. By linking predictions to meaningful interaction behaviors, educators and learners can better understand why certain performance outcomes are anticipated. This transparency is essential for the ethical and responsible deployment of learning analytics systems because it enables informed decision-making while reducing reliance on opaque algorithmic judgments.

These implications are particularly relevant for instructors and platform designers who require more than risk scores alone. If a system can identify whether poor outcomes are associated with fragmented navigation, declining engagement consistency, or weak self-regulation signals, interventions can be better targeted and more educationally meaningful. In this sense, the proposed framework supports not only prediction, but also diagnosis-informed educational design.

Supporting Explainable and Adaptive Learning Analytics:

Beyond performance prediction, the proposed HCI-informed approach supports explainable learning analytics by grounding model inputs in recognizable interaction behaviors. Rather than treating interaction data as abstract numerical features, this framework connects predictions to concrete aspects of user interaction, including navigation choices and engagement rhythms. This connection facilitates collaboration between data-driven systems and human stakeholders, including instructors and learning designers.

This interpretability also supports responsible use of predictive analytics in educational environments because stakeholders can understand and critically evaluate the behavioral evidence underlying model predictions.

Furthermore, the proposed framework can serve as a foundation for adaptive learning systems that respond to evolving learner behavior. By monitoring changes in interaction patterns over time, systems can adjust content presentation, feedback strategies, or interface elements to better align with individual learner needs. This adaptive capability reflects a transition from purely predictive analytics toward proactive and human-centered educational support.

Summary of HCI Contributions:

In summary, this study demonstrates that integrating human–computer interaction principles into learning behavior mining provides benefits at both predictive and interpretive levels. At the predictive level, interaction-aware features improve the ability to distinguish successful and at-risk learners. At the interpretive level, the same features provide a coherent behavioral framework through which interaction traces can be understood by educators, designers, and researchers. This combination strengthens the role of learning analytics as a tool for explanation, intervention, and human-centered educational improvement.

Recommendations for Educational Practice:

Based on the findings of this study, several recommendations can be proposed for educational practitioners, platform developers, and learning analytics researchers.

First, e-learning systems should incorporate interaction aware analytics rather than relying solely on aggregate activity counts. Features related to navigation consistency, temporal engagement regularity, and interaction diversity provide more meaningful insight into learner behavior and can support earlier identification of at-risk learners.

Second, predictive systems should emphasize interpretability and transparency to improve educational usability and stakeholder trust. Presenting instructors with understandable behavioral indicators may facilitate more informed intervention strategies compared with opaque prediction scores alone.

Third, adaptive e-learning platforms should integrate real-time monitoring of interaction patterns to support timely learner interventions. Behavioral signals such as declining engagement consistency or fragmented navigation may provide early evidence of disengagement and allow systems to respond proactively.

Finally, future educational analytics frameworks should prioritize human-centered design principles, ensuring that predictive technologies support learners and educators collaboratively rather than functioning as isolated automated systems.

Limitations and Future Work:

Limitations:

While this study provides valuable insights into human– computer interaction patterns for learning performance prediction, several limitations should be acknowledged.

First, the empirical evaluation was conducted using the Open University Learning Analytics Dataset, which, despite its richness and widespread use, represents a single platform context. Interface design, course organization, and learner demographics may differ substantially across e-learning systems, and these differences can influence observed interaction patterns. Consequently, the findings should be interpreted with caution when extending them to other platforms or learner populations.

Second, the proposed representation is derived from clickstream-based interaction logs recorded within the virtual learning environment. Although these logs are useful for capturing observable actions such as resource access and navigation behavior, they do not directly measure offline study behavior, learner motivation, attention, or cognitive load. The behavioral interpretations presented in this paper should therefore be understood as indirect inferences derived from digital interaction evidence rather than complete representations of the learner's internal cognitive state.

Third, while the proposed feature representation incorporates temporal aggregation, it does not explicitly model fine-grained sequential dependencies between individual interaction events. More expressive sequential models may capture additional behavioral nuances; however, they may also introduce challenges related to interpretability and computational complexity. Furthermore, extracting and maintaining fine-grained interaction features may become computationally demanding in large-scale learning environments involving substantial numbers of learners and interaction events. As learning platforms continue to expand, scalability and efficient feature extraction pipelines will become increasingly important considerations.

Although the proposed abstraction improves interpretability, it may simplify behaviors that are in practice influenced by multiple overlapping cognitive, motivational, and contextual factors. Real learner behavior is inherently multidimensional, and not all aspects of that complexity can be fully represented through interaction logs alone.

Finally, this study focuses primarily on learning performance prediction as the target outcome. Other important educational dimensions, including learning satisfaction, long-term knowledge retention, collaborative learning dynamics, and emotional engagement, were not considered and remain important directions for future investigation.

Future Work:

Future research can extend this work in several directions. One promising avenue involves evaluating the proposed HCI-informed learning behavior representation across multiple datasets and learning platforms to assess its generalizability. Comparative studies involving different interface designs, course structures, and instructional strategies could further clarify how human–computer interaction patterns vary across educational contexts.

In particular, early-stage prediction remains sensitive to sparse behavioral evidence, and future work should investigate uncertainty-aware approaches for low-data learning phases. Such methods may improve prediction reliability during the initial stages of learner interaction when available behavioral information is limited.

Another important direction involves integrating additional data modalities, including textual discussion content, eye-tracking signals, physiological measurements, or self-reported learner feedback, to enrich the interpretation of interaction behaviors. Combining interaction logs with complementary data sources may provide a more comprehensive understanding of learner engagement, cognition, and self-regulation processes.

Future work may also explore hybrid modeling approaches that combine interpretable feature-based representations with sequential or deep learning models. Such approaches may capture fine-grained temporal dependencies while maintaining transparency and human-centered interpretability.

Finally, the proposed framework could be extended toward real-time adaptive learning systems. By leveraging early detection of at-risk interaction patterns, future systems may deliver timely and personalized interventions that proactively support learners throughout the learning process. Investigating the educational effectiveness, ethical implications, and scalability of such adaptive interventions represents an important direction for future research.

Conclusion:

This study proposed an HCI-informed learning behavior representation for e-learning performance prediction using interaction log data from the Open University Learning Analytics Dataset (OULAD). Unlike conventional static feature aggregation approaches, the proposed framework incorporated navigation behavior, temporal engagement patterns, interaction diversity, and self-regulation signals to more effectively capture learner interaction dynamics.

Experimental evaluation demonstrated that HCI-informed features consistently improved predictive performance across logistic regression, random forest, and gradient boosting models. The best-performing configuration achieved an accuracy of 0.858, F1-score of 0.872, and AUC of 0.925, substantially outperforming traditional static baseline representations. Statistical validation further confirmed that these improvements were significant across all evaluated models.

Beyond predictive performance, the proposed framework contributes interpretable insights into how learners interact with digital learning environments. The findings demonstrate that learning analytics systems can become more explanatory, actionable, and human-centered when interaction structure is modeled explicitly rather than relying solely on aggregate activity volume. By connecting predictive outcomes with meaningful behavioral patterns, the proposed approach supports more transparent educational analytics and more informed intervention strategies.

Overall, this study highlights the value of integrating human–computer interaction principles into learning behavior mining for educational prediction tasks. Future research will extend the framework toward multi-platform validation, richer multimodal learning signals, uncertainty-aware prediction strategies, and adaptive intervention systems capable of supporting learners in real time.

References:

[1] G. Siemens, "Learning Analytics: The Emergence of a Discipline," *Am. Behav. Sci.*, vol. 57, no. 10, pp. 1380–1400, Oct. 2013, doi: 10.1177/0002764213498851.

[2] C. Romero and S. Ventura, "Educational data mining: A review of the state of the art," *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 40, no. 6, pp. 601–618, Nov. 2010, doi: 10.1109/TSMCC.2010.2053532.

[3] "(PDF) A Systematic Literature Review of Student' Performance Prediction Using Machine Learning Techniques." Accessed: May 15, 2026. [Online]. Available: https://www.researchgate.net/publication/354661323_A_Systematic_Literature_Review_of_Student'_Performance_Prediction_Using_Machine_Learning_Techniques

[4] "Predictive Analytics For Student Success: Ai-Driven Early Warning Systems And Intervention Strategies For Educational Risk Management," *Educ. Res. Hum. Dev.*, vol. 2, no. 2, 2025, doi: 10.61784/erhd3042.

[5] S. Chidambaranathan, B. Puli, P. Sundaramoorthy, N. N. Jose, R. Praveen, and R. Daruvuri, "INSIGHT: A next-generation framework for proactive mental health detection using advanced data fusion and deep learning," *Int. J. Sci. Res. Arch.*, vol. 14, no. 2, pp. 821–827, Feb. 2025, doi: 10.30574/IJSRA.2025.14.2.0453.

[6] "View Learning Analytics: Challenges and Future Research Directions | eleed - e-learning and education." Accessed: May 15, 2026. [Online]. Available: https://www.eleed.de/eleed/article/view/138/230

[7] Hassan Khosravi, Simon Buckingham Shum, "Explainable Artificial Intelligence in education," *Comput. Educ. Artif. Intell.*, vol. 3, p. 100074, 2022, doi: https://doi.org/10.1016/j.caeai.2022.100074.

[8] Y. Dimitriadis, R. Martínez-Maldonado, and K. Wiley, "Human-Centered Design Principles for Actionable Learning Analytics," *Res. E-Learning ICT Educ. Technol.*

Pedagog. Instr. Perspect., pp. 277–296, Jan. 2021, doi: 10.1007/978-3-030-64363-8_15/SAVE-RESEARCH.

[9] Riordan Alfredo, Vanessa Echeverria, "Human-centred learning analytics and AI in education: A systematic literature review," *Comput. Educ. Artif. Intell.*, vol. 6, p. 100215, 2024, doi: https://doi.org/10.1016/j.caeai.2024.100215.

[10] Kenneth Holstein, Jennifer Wortman Vaughan, Hal Daumé III, Miro Dudík, Hanna Wallach, "Improving fairness in machine learning systems: What do industry practitioners need?," *arXiv:1812.05239*, 2019, [Online]. Available: https://arxiv.org/abs/1812.05239

[11] Carmen Fernández-Morante, Beatriz Cebreiro-López, "Adaptive Learning Supported by Learning Analytics for Student Teachers' Personalized Training during in-School Practices," *Sustainability*, vol. 14, no. 1, p. 124, 2021, doi: 10.3390/su14010124.

[12] "(PDF) Predicting students' performance in distance learning using machine learning techniques." Accessed: Mar. 19, 2026. [Online]. Available: https://www.researchgate.net/publication/233291686_Predicting_students'_performance_in_distance_learning_using_machine_learning_techniques

[13] Hassam Iqbal, Jinchuan Zheng, Rifai Chai & Sivachandran Chandrasekaran, "Electric powered wheelchair control using user-independent classification methods based on surface electromyography signals," *Med. Biol. Eng. Comput.*, vol. 62, pp. 167–182, 2024, [Online]. Available: https://link.springer.com/article/10.1007/s11517-023-02921-z

[14] "(PDF) Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence." Accessed: May 15, 2026. [Online]. Available: https://www.researchgate.net/publication/267510046_Learning_Analytics_and_Educational_Data_Mining_in_Practice_A_Systematic_Literature_Review_of_Empirical_Evidence

[15] T. Rajesh Kumar, T. Vamsidhar, B. Harika, T. Madan Kumar, and R. Nissy, "Students performance prediction using data mining techniques," *Proc. Int. Conf. Intell. Sustain. Syst. ICISS 2019*, pp. 407–411, Feb. 2019, doi: 10.1109/ISS1.2019.8907945.

[16] D. A. Shafiq, M. Marjani, R. A. A. Habeeb, and D. Asirvatham, "Student Retention Using Educational Data Mining and Predictive Analytics: A Systematic Literature Review," *IEEE Access*, vol. 10, no. July, pp. 72480–72503, 2022, doi: 10.1109/ACCESS.2022.3188767.

[17] Feiyue Qiu, Lijia Zhu, "E-Learning Performance Prediction: Mining the Feature Space of Effective Learning Behavior," *Entropy*, vol. 24, no. 5, p. 722, 2022, doi: https://doi.org/10.3390/e24050722.

[18] Zara Ersozlu, Sona Taheri & Inge Koch, "A review of machine learning methods used for educational data," *Educ. Inf. Technol.*, vol. 29, pp. 22125–22145, 2024, [Online]. Available: https://link.springer.com/article/10.1007/s10639-024-12704-0

[19] Tumaini Kabudi, Ilias Pappas, "AI-enabled adaptive learning systems: A systematic mapping of the literature," *Comput. Educ. Artif. Intell.*, vol. 2, p. 100017, 2021, doi: https://doi.org/10.1016/j.caeai.2021.100017.

[20] P. Leitner, M. Khalil, and M. Ebner, "Learning analytics in higher education—a literature review," *Stud. Syst. Decis. Control*, vol. 94, pp. 1–23, 2017, doi: 10.1007/978-3-319-52977-6_1.

[21] Wannisa Matcha, Dragan Gašević, "Analytics of Learning Strategies: Associations with Academic Performance and Feedback," *ACM Int. Conf. Proceeding Ser.*, 2019, [Online]. Available: https://www.researchgate.net/publication/331337849_Analytics_of_Learning_Strategies_Associations_with_Academic_Performance_and_Feedback

[22] Susanne de Mooij, Joni Lämsä, Lyn Lim, Olli Aksela, Shruti Athavale, Inti Bistolfi, Flora Jin, Tongguang Li, "A Systematic Review of Self-Regulated Learning through Integration of Multimodal Data and Artificial Intelligence," *Educ. Psychol. Rev.*, vol. 37, no. 54, 2025, [Online]. Available: https://link.springer.com/article/10.1007/s10648-025-10028-0

[23] Abdalganiy Wakjira, Samit Bhattacharya, "Predicting Student Engagement in the Online Learning Environment," *Int. J. Web-Based Learn. Teach. Technol.*, vol. 16, no. 6, pp. 1–21, 2021, doi: 10.4018/IJWLTT.287095.

[24] R. S. Baker and P. S. Inventado, "Educational Data Mining and Learning Analytics," *Learn. Anal. From Res. to Pract.*, pp. 61–75, Jan. 2014, doi: 10.1007/978-1-4614-3305-7_4.

[25] P. Dillenbourg and P. Jermann, "Technology for Classroom Orchestration," *New Sci. Learn. Cogn. Comput. Collab. Educ.*, pp. 525–552, 2010, doi: 10.1007/978-1-4419-5716-0_26.

[26] A. Pardo and G. Siemens, "Ethical and privacy principles for learning analytics," *Br. J. Educ. Technol.*, vol. 45, no. 3, pp. 438–450, May 2014, doi: 10.1111/bjet.12152.

[27] "Interpretable Machine Learning." Accessed: May 15, 2026. [Online]. Available: https://christophm.github.io/interpretable-ml-book/

[28] "Open University Learning Analytics Dataset (OULAD)." Accessed: Mar. 19, 2026. [Online]. Available: https://www.kaggle.com/datasets/anlgrbz/student-demographics-online-education-dataoulad