Cómo referenciar este artículo / How to Reference this Article
Jiang, Z., Karim, A. A., & Khalid, F. (2025). Machine Learning–Based Assessment of Student Engagement in Blended Learning: A Case Study in Chinese Higher Education. *Foro de Educación*, 23(2), 111-131. doi: https://doi. org/10.14201/fde.23206

Machine Learning–Based Assessment of Student Engagement in Blended Learning: A Case Study in Chinese Higher Education

Evaluación del compromiso estudiantil en el aprendizaje combinado mediante aprendizaje automático: un estudio de caso en la educación superior china

Zhenfeng Jiang*
Faculty of Education, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia.
School of Information Science and Engineering, Zaozhuang University, 277160 Zaozhuang, China.
Email: p126242@siswa.ukm.edu.my

Aidah Abdul Karim
Faculty of Education, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia.
Email: eda@ukm.edu.my

Fariza Khalid
Faculty of Education, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia.
Email: fariza.khalid@ukm.edu.my

Recibido / Received: 09/05/2025
Aceptado / Accepted: 17/08/2025

Abstract: This study investigates student involvement in a blended learning course within Chinese higher education and assesses the potential of machine learning to facilitate ongoing engagement evaluation across time. Engagement is conceptualised within a dual-subject framework that examines the dynamic interplay between learners and instructors, and is implemented across behavioural, cognitive, and social-interaction dimensions. Data from 144 students over three semesters in an undergraduate course on Database Principles and Applications produced eleven indicators based on Learning Management Systems (LMS) activity logs and classroom behaviours. Predictive models were created utilising both batch learning and an incremental Random Forest, with "incremental" indicating a model that updates with fresh data without necessitating complete retraining. K-Means clustering was subsequently utilised to delineate unique engagement characteristics. The results indicate that the incremental model attained superior accuracy and demonstrated enhanced adaptability to newly arriving data. Clustering demonstrated diverse engagement patterns that underscore the importance of varied educational methodologies. Longitudinal observations demonstrated that modifications in instructional design positively influenced student engagement. This study enhances current LMS-based engagement research by integrating the dual-subject paradigm with progressively updated learning analytics. It also illustrates a pragmatic approach for educators to consistently assess engagement and implement timely, data-driven modifications in blended higher education settings.

Keywords: Blended Learning, Machine Learning-Based Evaluation Models, Learning Behavior Data Analysis, Student Engagement Assessment.

1. Introduction

Blended learning, integrating in-person instruction with online activities, has emerged as a prevalent method in global higher education. This paradigm offers flexibility and tailored learning opportunities, enabling institutions to more effectively address various learner requirements and foster a student-centered educational approach (Boelens, De Wever & Voet, 2017; Pramesworo et al., 2023). Chinese colleges have rapidly integrated technology-enhanced tools, such as Learning Management Systems (LMS), mobile applications, and digital resources, transforming classrooms into interactive and resource-rich environments. Notwithstanding these advancements, enduring difficulties persist. A considerable number of students exhibit minimal initiative, diminished engagement, and inadequate collaboration in blended courses, hence undermining the overall efficacy of this instructional paradigm (Pramesworo et al., 2023; Sareen & Mandal, 2024).

Student engagement is a multifaceted concept that significantly impacts academic performance, cognitive development, motivation, and retention (Islam, Sarker & Islam, 2022; Wong et al., 2024). This study operationalises engagement through behavioural, cognitive, and social-interaction dimensions, encompassing students' behaviours, cognitive investment, and interactions with peers and instructors. Although extensive research has focused on involvement in face-to-face or online contexts, there is less understanding of its expression in mixed environments, particularly within Chinese higher education, where cultural, institutional, and technical aspects uniquely converge (Yu et al., 2025). This constraint restricts the development of interventions capable of adapting to both online and in-person learning environments.

Current engagement evaluation techniques, including questionnaires and psychometric measures, are frequently employed owing to their practicality. Nonetheless, these instruments frequently depend on self-reported data, which may be affected by biases, issues of reliability, and restricted causal inference (Manwaring et al., 2017; Zhoc et al., 2019). Furthermore, they generally offer only static representations of engagement instead of ongoing insights, rendering them inadequate for dynamic mixed settings where learning behaviours change over time (Huang, Kuang & Ling, 2022). These problems highlight the necessity for data-driven, process-oriented methodologies that can encapsulate the multifaceted and dynamic essence of student participation.

To address this requirement, educational data mining (EDM) and learning analytics (LA) have been utilised for engagement evaluation. Previous research has utilised Learning Management Systems or other digital footprints; nevertheless, the majority emphasise static metrics or singular engagement variables and seldom include continuous instructional interventions. Aldowah, Al-Samarraie and Fauzy (2019) utilised LMS data to forecast engagement in distance learning contexts, whereas Salas-Pilco, Yang and Zhang (2022) performed a systematic review of behavioural, cognitive, and affective engagement in Latin American higher education amid the COVID-19 pandemic. Yılmaz and Yılmaz (2023) examined the impact of sociability, community sense, and course satisfaction on involvement in flipped classrooms facilitated by virtual learning communities. These studies illustrate the significance of learning analytics; nonetheless, deficiencies persist in the dynamic evaluation of involvement in blended courses alongside the incorporation of educators' instructional modifications.

In blended learning, students engage in both online and face-to-face activities, with their educational experiences influenced by technological platforms and the instructional methods utilised by educators. Pedagogical interventions—such as modifications to content delivery, instructional tempo, and feedback systems—can profoundly affect students' motivation and engagement. Nevertheless, the majority of current research neglects the continuous influence of instructional refinement on student involvement, leading to static or inadequate evaluation frameworks. This study tackles existing constraints by presenting a machine learning-based engagement evaluation model that incorporates incremental learning approaches alongside multidimensional engagement indicators, while also considering a dual-subject perspective. The methodology seeks to deliver a more precise, adaptable, and actionable assessment in blended higher education courses by monitoring student activity across several semesters.

Consequently, the study examines the subsequent research enquiries:

RQ1: In what ways may behavioural data obtained from Learning Management Systems (LMS) be employed to derive multidimensional indices of student engagement within a blended learning context?

RQ2: How can student engagement levels be accurately categorised, taking into account the impact of ongoing educational enhancements by educators?

RQ3: What methodology can be employed to create a machine learning model for evaluating student involvement in blended learning through the specified engagement indicators?

This research conceptually enhances the comprehension of dynamic, multidimensional student participation by addressing these questions. It effectively promotes the creation of adaptive, data-driven evaluation instruments that enable ongoing educational enhancement. This project seeks to cultivate a more adaptive educational environment that encourages active student engagement and facilitates ongoing instructional innovation in higher education.

2. Related Works

2.1. Blended Learning Based on the Dual-Subject Model

Blended learning, which combines in-person instruction with online educational experiences, has become a prevalent concept in higher education. It seeks to utilise the advantages of both conventional and digital teaching methods to facilitate adaptable, learner-focused, and inclusive educational settings (Islam et al., 2022; Pramesworo et al., 2023). This study investigates engagement and learning interactions using the dual-subject paradigm, which highlights the active responsibilities of both teachers and students. This viewpoint is pertinent to our research, as it underscores the collaborative influence of teacher interventions and student behaviours on engagement, a concept implemented in our machine learning (ML) architecture. In this paradigm, educators function not merely as conveyors of knowledge but also as facilitators, mentors, and architects of learning experiences. Their duties encompass the coordination of educational activities during several temporal stages: pre-class preparation, in-class engagement, and post-class reflection. Simultaneously, educators must be cognisant of students' prior knowledge, digital literacy, and emotional preparedness—particularly in blended environments where support systems are inconsistent. In contrast, students are anticipated to participate cognitively and behaviourally through independent learning, self-evaluation, and collaborative investigation, hence promoting more profound and lasting learning results (Boelens et al., 2017; McGee & Reis, 2012; Sareen & Mandal, 2024).

Instructional strategies in blended education differ based on learning objectives. Receptive methods, including lectures and multimedia presentations, are typically employed for factual knowledge, often enhanced by digital modules for repeated access and review (Nguyen, 2017). In contrast, participatory strategies—such as flipped classrooms, case-based discussions, simulations, and group projects—are favoured for procedural and conceptual understanding. These methodologies encourage active investigation, discussion, and the application of information in genuine circumstances (García-Peñalvo, Fidalgo-Blanco & Sein-Echaluce, 2018; Liu et al., 2024). To ensure effectiveness in these designs, the learning environment must be meticulously organised to support varied learners, especially those who may lack familiarity with collaborative or self-directed methodologies. Consequently, an inclusive mixed environment necessitates both educational innovation and a focus on fairness regarding technological access, workload distribution, and feedback literacy (Sareen & Mandal, 2024).

The dual-subject model underscores constructive alignment, which involves the deliberate connection of learning outcomes, instructional activities, and assessment strategies to guarantee coherence and learner accountability during the blended learning process (McGee & Reis, 2012; Sareen & Mandal, 2024). This technique, when executed proficiently, augments motivation, facilitates differentiated instruction, and cultivates learners' potential for lifelong learning. Furthermore, by allocating agency to both educators and learners, the approach advances the overarching objectives of participatory and democratic education.

2.2. Student Engagement in Blended Learning Environments

Student involvement is acknowledged as a multifaceted construct that profoundly affects academic achievement, course satisfaction, and general well-being in higher education (Wong et al., 2024; Zhoc et al., 2019). In blended learning contexts, engagement assumes novel forms and problems owing to the hybrid character of learning activities and communication channels. Engagement is not merely an individual cognitive condition; it also reflects learners' interactions with their environment, encompassing social, technological, and institutional frameworks that facilitate or hinder their participation. This ecological perspective on involvement has garnered heightened attention in culturally varied settings, especially where educational inequalities endure.

Utilising the tripartite framework established by Fredricks, Blumenfeld and Paris (2004), involvement is often characterised by behavioural, cognitive, and emotional dimensions. In blended learning, social interaction is a significant and essential aspect, intricately linked to motivation and the feeling of belonging within learning communities (Saichaie, 2020; Sinha et al., 2015). This study delineates involvement through three dimensions—behavioral, cognitive, and social interaction—modifying the traditional paradigm to accommodate the mixed setting in which communication with peers and instructors is paramount.

2.2.1. Behavioral Engagement.

Behavioural engagement denotes visible activities that reflect a student's involvement in learning, including attendance, participation, and task fulfilment. In blended environments, it includes online activities (e.g., viewing videos, participating in forums) and offline actions (e.g., group discussions, classroom exercises). The integration of synchronous and asynchronous elements facilitates involvement by providing students with flexibility in their engagement timing and methods (Manwaring et al., 2017). Data obtained from LMS platforms and digital logs yield significant insights into behaviours and enable teachers to track engagement trends. Furthermore, behavioural data might reveal nuanced discrepancies; for instance, students from under-resourced families may engage differently due to bandwidth limitations, caregiving obligations, or insufficient private study environments, necessitating more adaptive instructional design.

2.2.2. Cognitive Engagement

Cognitive engagement refers to learners' commitment to mastering knowledge by employing deep learning methodologies, critical thinking, and self-regulation. Jovanović et al. (2017) emphasise the significance of flipped classrooms in enhancing metacognitive awareness and strategic learning behaviours. Blended learning redistributes classroom time to prioritise higher-order cognitive skills, including analysis, assessment, and synthesis, while enabling independent acquisition of core knowledge via digital resources (Liu et al., 2024; McGee & Reis, 2012). To effectively engage learners intellectually, instructors must take into account previous experience to self-directed learning paradigms. In areas where conventional lecture-based instruction prevails, facilitating students' transition to active participation necessitates scaffolding, reflective assignments, and formative feedback customised to diverse cognitive preparation levels.

2.2.3. Social Interaction Engagement.

Social connection is essential for cultivating connectedness, trust, and peer learning. It encompasses both official and informal interactions among students and between students and teachers, encouraged by classroom discussions, collaborative activities, and digital communication platforms. Blended settings provide enhanced potential for meaningful and synchronous interpersonal encounters compared to fully online alternatives (Saichaie, 2020). Yılmaz and Yılmaz (2023) underscore the significance of creating social environments—such as group chats, surveys, and discussion boards—that promote inclusive engagement and input. These connections are fundamental to establishing learning communities and fostering persistence, particularly among less confident learners.

Moreover, feedback and self-reflection mechanisms—such as peer review, formative assessments, and reflection journals—facilitate students' processing of their learning experiences and link emotional engagement with academic objectives (Fredricks et al., 2004; Salas-Pilco et al., 2022). In culturally diverse classrooms, inclusive social engagement necessitates the consideration of linguistic preferences, varying comfort levels with peer critique, and differing expectations concerning teacher-student power dynamics—all of which must be meticulously addressed to establish equitable blended learning experiences.

2.3. Application of Machine Learning in Learning Behavior Analysis

This research utilises incremental machine learning models for the continuous measurement of engagement, enabling the ongoing monitoring of student behaviours and the adaptation of engagement levels in response to instructional interventions over the course of a semester. The emergence of digital learning platforms, especially Learning Management Systems (LMSs), has empowered educators to gather and scrutinise extensive behavioural data in both online and offline educational environments. The data—encompassing login frequency, resource access, quiz attempts, and discussion patterns—constitute the foundation for learning analytics, a burgeoning discipline focused on comprehending and enhancing learning processes (Aldowah et al., 2019; Almusaed et al., 2023). In fact, these data frequently represent the instructional design of a course and the wider institutional context, necessitating the interpretation of behaviour not as immutable learner characteristics but as reactions to particular learning contexts.

Machine learning (ML), a branch of artificial intelligence, provides sophisticated techniques for revealing concealed patterns, forecasting academic results, and customising training. In contrast to conventional statistical methods, machine learning algorithms, including support vector machines, decision trees, neural networks, and ensemble techniques such as random forests, are superior in modelling intricate and nonlinear relationships (García-Peñalvo et al., 2018; Hung et al., 2020). For instance, machine learning algorithms can identify preliminary indicators of disengagement, categorise learners into risk classifications, or suggest customised actions to enhance educational results.

This study specifically examines incremental learning, which continuously updates as new data is received without need complete retraining. By incorporating incremental learning with multidimensional engagement metrics, the model adaptively monitors students' learning trajectories across time. This method is particularly pertinent in blended courses because student behaviours change and educational tactics may adapt. Incremental learning allows the model to maintain adaptability across a semester, considering elements such as previous preparation, technical proficiency, and cultural learning preferences (Almusaed et al., 2023; Jovanović et al., 2017).

Recent studies illustrate the significance of machine learning-driven engagement analytics for real-time monitoring. Integrating behavioural analytics into course design can enhance student performance and attitudes towards learning (Yu et al., 2025). As educational institutions adopt data-informed teaching methods, machine learning-driven engagement analytics present a promising avenue for adaptive, scalable, and equitable learning experiences, particularly in large-enrollment or resource-constrained environments where manual engagement tracking is unfeasible. However, these methodologies must be implemented judiciously, guaranteeing that algorithmic forecasts do not perpetuate stereotypes or exacerbate existing disparities among learners. Instead, machine learning should function as a mechanism for recognising systemic obstacles and formulating more inclusive educational trajectories, consistent with the social justice principles advocated in critical pedagogical traditions (Adeleye, Eden & Adeniyi, 2024).

3. Materials and Methods

3.1. Research Design

This project employs a quantitative design that combines learning analytics with machine learning techniques to analyse student engagement in a dual-subject blended teaching strategy. The design is based on the premise that student involvement is influenced by both individual behaviours and the educational intents and adaptive tactics utilised by instructors during the course (Islam et al., 2022).

Blended education settings are intrinsically dynamic, necessitating instructors to consistently modify instructional tasks, resource distribution, and classroom activities in response to technology capabilities, learner feedback, and previous teaching experiences (Boelens et al., 2017). The study highlights the necessity of real-time and longitudinal data gathering over three semesters, facilitating a process-oriented analysis of changing interaction patterns over time. This work utilises machine learning techniques based on educational data mining (EDM) to tackle the complex aspects of engagement, encompassing behavioural, cognitive, and social interaction dimensions. This methodology transcends descriptive research, facilitating the identification of concealed patterns and predicting correlations between instructional design elements and engagement metrics (Aldowah et al., 2019; Hung et al., 2020). In accordance with recent advances in blended learning research (Yu et al., 2025), the application of machine learning facilitates more sophisticated assessments of the effects of educational interventions on learning trajectories across cohorts.

3.2. Sample

The study sample comprised 249 second-year undergraduate students majoring in computer science at a public university in southern Shandong Province, China. Participants were enrolled in the foundational course Database Principles and Applications during three successive semesters from 2022 to 2024. The allocation of students by semester was as follows: 101 in Semester 1, 99 in Semester 2, and 49 in Semester 3.

A cluster sampling strategy was utilised, selecting intact class groupings to maintain ecological validity and observe naturally occurring learning behaviours. This method facilitated the acquisition of genuine engagement data inside standard classroom dynamics. All research protocols complied with ethical standards: informed consent was secured from all participants, and personal identifiers were omitted during data processing. The study protocol obtained ethical approval from the Institutional Review Board (IRB) of Zaozhuang University.

3.3. Research Setting

The research was performed in the blended learning setting of the Database Principles and Applications course, an essential element of the computer science curriculum. The curriculum is organised into three interconnected domains: (i) theoretical knowledge, (ii) practical application, and (iii) project-based problem-solving. Every domain is presented via task-oriented learning modules, aligning with blended learning's focus on active and independent participation (Pramesworo et al., 2023; Saichaie, 2020).

Each task module consists of three pedagogically interconnected stages:

Self-directed learning prior to class: Students interact with essential knowledge through the LMS (Rain Classroom), encompassing video lectures, interactive exercises, and diagnostic quizzes. These resources facilitate the activation of prior information and encourage self-directed learning (Nguyen, 2017).

In-class collaborative learning: In-person courses prioritise peer connection and immediate involvement, with case discussions, polling exercises, group assignments, and instructor-led elucidation of intricate subjects. Mobile applications and LMS technologies were utilised to gather student responses and deliver prompt feedback, enhancing engagement and formative evaluation (Manwaring et al., 2017).

Reflection and assignments following class: Students undertake solitary assignments including programming exercises, assessments, and reflective compositions. Submissions are made via the LMS, which monitors submission timestamps, performance data, and participation records. This data is essential for tracking ongoing involvement over time.

The LMS platform utilised in this study (Rain Classroom) provides comprehensive data collecting functionalities, encompassing attendance records, video interaction logs, quiz results, discussion forum engagement, and assignment submissions. The Classroom Activity Management module allows teachers to implement real-time engagement assignments (e.g., in-class quizzes, group polls), providing immediate insights into student participation and understanding. This multi-phase educational paradigm conforms with best practices in blended pedagogy and enables comprehensive process-level data collecting. Furthermore, it facilitates customised instruction, allowing educators to customise learning activities based on students' involvement levels, performance records, and educational requirements (Sinha et al., 2015; Wong et al., 2024).

3.4. Research Instrument

3.4.1. Learning Activities Representing Blended Learning Engagement

This research employs a multidimensional paradigm for student engagement, based on the seminal work of Fredricks et al. (2004) and subsequent modifications by Zhoc et al. (2019). Engagement is defined through three dimensions: behavioural, cognitive, and social interaction. Data encompassing these dimensions is obtained from digital traces logged in the Learning Management System (LMS) and functions as essential inputs for subsequent modelling. Behavioural engagement denotes students' adherence to course obligations and involvement in educational activities. It includes LMS login frequency, task completion rates, assignment submissions, quiz participation, and engagement during synchronous sessions (Islam et al., 2022; Manwaring et al., 2017).

Cognitive engagement denotes the extent of students' mental exertion, encompassing their use of self-regulated learning techniques and reflective methodologies. Indicators comprise the frequency of video segment rewatching, the quality of assignments, performance on in-class case tasks, and overall test scores (Hung et al., 2020; Sareen & Mandal, 2024). Social interaction engagement assesses students' communication and collaboration with classmates and instructors, which is vital for enhancing motivation and building a sense of academic belonging. This is measured by the quantity of discussion forum posts, comments, and involvement in peer reviews (Wong et al., 2024; Yılmaz & Yılmaz, 2023). This approach facilitates thorough monitoring of engagement during various instructional phases—pre-class, in-class, and post-class—offering an in-depth, process-oriented perspective on learner involvement.

Foro de Educación, v. 23, n. 2, julio-diciembre / july-december 2025, pp. 111-131.
e-ISSN: 1698-7802

3.4.2. Assessment Indicator Variables

Eleven distinct activity-based indicators were delineated from the LMS data, categorised within the three aspects of engagement. These indicators encompass both direct and indirect expressions of involvement, providing a multifaceted viewpoint to guide the development of machine learning models. Table 1 delineates the operational definitions and contextualises each indicator within pertinent learning activities and instructional phases.

Table 1: Assessment Indicators for Student Engagement in Blended Learning.

Dimension	Indicator Variable	Learning Activity Description	Instructional Phase
Behavioral Engagement	task_p learn_count assignment_p test_ participation classroom_ interaction	Percentage of course tasks completed via LMS. Frequency of LMS access for self-learning. Percentage of submitted assignments. Completion rate of quizzes and chapter tests. Performance in interactive activities (e.g., questioning, group discussion, voting).	Pre-class, In-class, Post-class Post- class Pre-class, In-class, Post-class In-class, Post-class In-class
Cognitive Engagement	video_p assignment_ grade class_task_ grade test_grade	Percentage of video segments with repeat viewing (>100%), indicating reflective engagement. Quality of submitted assignments. Scores on classroom case-based tasks. Overall test performance throughout the course.	Pre-class Post- class In-class Pre-class, In-class, Post-class
Social Interaction Engagement	discuss_sum mutual evaluation	Number of forum posts and comments. Participation rate in peer review activities.	Pre-class, In-class, Post-class Post-class

3.5. Data Analysis

The data analysis procedure unfolded in three primary stages:

Phase 1: Data preprocessing. The primary objective of this initial step was to enhance data quality and ready the dataset for substantive analysis. Common data concerns, including absent values, discrepancies, and noise, were rectified through cleaning, integration, and standardisation. Transformations and discretisation techniques were employed when necessary to ensure the dataset's structure conformed to the requirements of machine learning algorithms, hence improving the validity and reliability of subsequent studies.

Phase 2: Creation of engagement labels. Due to the intricacy and magnitude of behavioural data collected by the LMS, human categorisation of engagement levels proved impractical and subjective. To tackle this issue, unsupervised clustering methods, such as K-Means and hierarchical clustering, were utilised for their computational efficiency, interpretability, and effectiveness in identifying latent engagement patterns within a moderately sized dataset. The resultant clusters functioned as proxy engagement labels, offering a data-driven and scalable substitute for supervised model training. Although alternative clustering methods (e.g., Gaussian Mixture Model, spectral clustering) may effectively identify more intricate structures, they were not utilised in this study to ensure methodological transparency and compatibility with later supervised modelling. This method enabled the recognition of underlying behavioural patterns indicative of differing levels of cognitive and emotional involvement (Salas-Pilco et al., 2022).

Phase 3: Development of an evaluation framework. A model for evaluating learner engagement was created utilising machine learning algorithms, such as Decision Trees, Random Forests, and Support Vector Machines (SVM). Random Forest and SVM were chosen for their resilience in managing high-dimensional data, capacity to model nonlinear correlations, and, specifically for Random Forest, suitability for incremental learning in longitudinal tracking. Both non-incremental and incremental model construction methodologies were examined to enhance predictive accuracy. The execution of this dual technique enabled the identification of an effective evaluation framework, assuring the adaptability and robustness of the assessment model in various learning situations. Although deep learning and hybrid ensemble methods may provide superior predictive performance, they are data-intensive and less interpretable. The dataset, consisting of merely 249 students over three semesters, was too inadequate to adequately support such methodologies without the risk of overfitting and jeopardising generalisability. Consequently, these advanced techniques were identified as avenues for future research instead of being utilised in the present study. To investigate the efficacy and flexibility of the proposed engagement evaluation approach, numerous comparison experiments were performed utilising both batch and incremental learning frameworks. The subsequent part delineates the experimental results, emphasising the model's predictive efficacy across various algorithms and learning methodologies, as well as its ability to monitor longitudinal variations in student involvement.

4. Research Findings

4.1. Data Preprocessing

Thorough data preparation was essential for guaranteeing the validity and robustness of the engagement assessment model. The unprocessed dataset exhibited two principal challenges: absent values and erratic data. A regression-based imputation technique was employed to address missing variables. This method approximated missing data points by utilising correlations among existing characteristics, maintaining the dataset's internal structure without altering variable distributions. Records with numerous missing characteristics were eliminated by a tuple elimination procedure to ensure the integrity of the analytical sample. Cluster analysis revealed the presence of noisy data, including extreme outliers. Observations situated at considerable distances from their corresponding cluster centroids were either rectified—when discernible through logical patterns—or eliminated outright if considered invalid. Supplementary preparation measures encompassed deduplication to remove redundant records, z-score normalisation to harmonise feature scales, and consistency checks to guarantee standardised data entry. The thorough preparation procedure produced a clean, dependable dataset that formed the basis for the next clustering and model construction phases. These measures align with optimal practices in educational data mining for blended learning contexts (Almusaed et al., 2023; Hung et al., 2020).

4.2. Learner Engagement Label Acquisition

4.2.1. Clustering Scheme Identification

Patterns of learner involvement were discerned by K-Means clustering over three semesters (n=249). Pedagogical adjustments between semesters led to varied student behaviours. One-way ANOVA analyses were performed on all 11 engagement measures to evaluate inter-semester variability. Ten out of eleven variables exhibited significant variations across semesters ($p < 0.05$), underscoring the need for distinct clustering for each semester (refer to Appendix A, Table A1).

4.2.2. Data Standardization

A Pearson correlation analysis was performed for each semester to identify multicollinearity among engagement factors, hence optimising clustering performance. Indicator pairs with correlation coefficients beyond 0.8 were examined for redundancy. Variables with high correlation were deliberately eliminated to avoid grouping distortion and enhance interpretability (García-Peñalvo et al., 2018).

Subsequent to feature selection, the dataset for each semester was independently standardised by z-score normalisation. This technique ensured the comparability of feature scales and mitigated potential biases arising from differences in measurement ranges across different time periods and student cohorts. Figures 1 and 2 depict the inter-variable correlation matrix and the final collection of standardised indicators employed in clustering.

This curated and standardised dataset facilitated the creation of dependable engagement labels, crucial for the supervised machine learning models devised in the subsequent research phase.

Figure 1: Heat Map of Indicator Correlation for the Three Semesters Dataset.

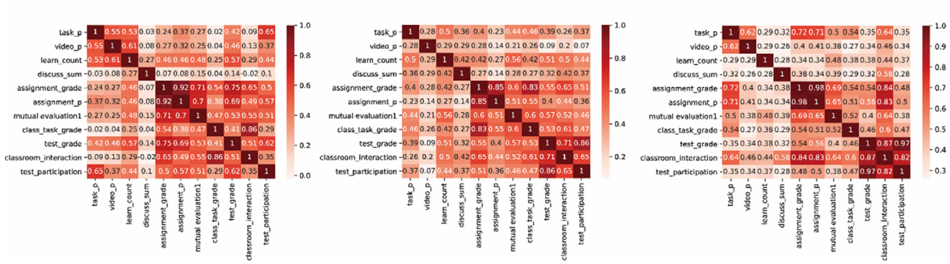
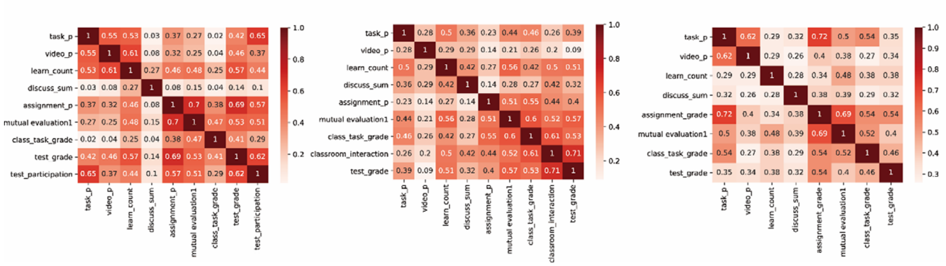


Figure 2: Heat Map of Indicator Relevance for Clustering Algorithms.



4.2.3. Cluster Quality Inspection and Result Analysis

The efficacy of the K-Means clustering algorithm predominantly hinges on identifying an ideal number of clusters (k). To resolve this, both the Elbow Method and the Silhouette Coefficient were utilised. The findings from these two complementary methodologies consistently indicated that three clusters (k = 3) achieved the optimal equilibrium between model simplicity and explanatory power.

To enhance the validation of the clustering results, one-way ANOVA tests were performed on the 11 engagement markers for each semester. ANOVA revealed significant differences among clusters for the majority of indicators, with the exception of discuss_sum in Semester 1 (p = 0.4025; refer to Appendix A, Table A2).

Figure 3 displays a radar plot to visualise engagement profiles across clusters, highlighting the relative strengths and weaknesses of each engagement group. Due to the variability of engagement markers across semesters, cluster label interpretations were standardised to low, moderate, and high engagement for uniformity.

Figure 3: Radar Plot of Clustering Results for the Three-semester Data Sets.

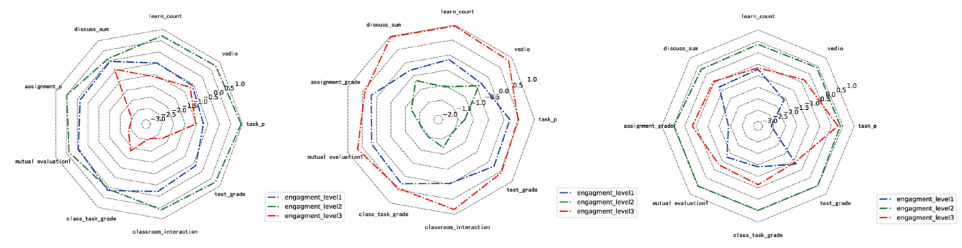


Table 2 clearly illustrates that Cluster 2 (high engagement) students consistently demonstrated superior performance across nearly all indicators, particularly in task completion, video learning, discussion participation, and assignment quality. Learners in Cluster 0 (moderate engagement) exhibited intermediate behaviors, while Cluster 1 (low engagement) showed the lowest levels of engagement across all dimensions.

Table 2: Means Statistics of 3 Clustering Indicator Values for 3 Semesters.

Engagement Level	Semester	Task_p	Video_p	Learn_Count	Discuss_Sum	Assignment_Grade	Assignment_p	Mutual Evaluation	Class_Task_Grade	Test_Grade	Classroom_Interaction	Test_Participation
mid_level (Cluster 0)	term1	0.19	0.04	139.73	1.04	83.75	0.90	0.76	86.07	77.62	85.39	0.73
	term2	0.93	0.15	263.15	5.91	86.60	0.96	0.74	88.85	72.12	72.59	0.86
	term3	0.97	0.57	141.09	7.27	85.55	0.86	0.44	41.00	61.50	57.35	0.73
low_level (Cluster 1)	term1	0.03	0.01	65.50	0.00	50.01	0.58	0.32	61.45	44.23	59.17	0.41
	term2	0.45	0.11	135.21	2.05	68.79	0.82	0.29	65.84	41.13	58.64	0.60
	term3	0.31	0.20	133.83	2.17	63.98	0.63	0.31	17.50	63.25	46.41	0.73
high_level (Cluster 2)	term1	0.87	0.36	286.49	1.49	87.90	0.98	0.84	85.02	89.05	86.96	0.92
	term2	1.00	0.44	415.81	17.85	91.10	0.99	0.88	90.63	77.54	82.92	0.91
	term3	1.00	0.80	258.19	18.44	96.42	0.97	0.80	74.75	88.86	73.27	

4.2.4. Correlation Analysis between Engagement Patterns and Academic Performance

A correlation analysis was performed to evaluate the pedagogical significance of

Foro de Educación, v. 23, n. 2, julio-diciembre / july-december 2025, pp. 111-131.
e-ISSN: 1698-7802

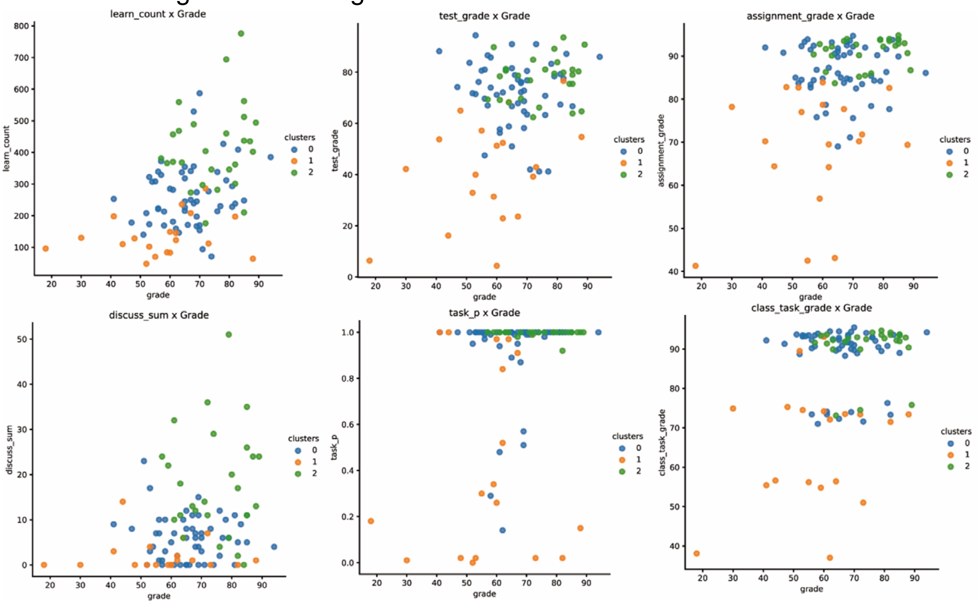
engagement-based clustering by examining the relationship between engagement levels and academic success, as indicated by final course grades. Previous studies emphasise that characteristics of engagement must significantly link with academic performance to be considered credible markers of learning quality (García-Peñalvo et al., 2018).

Figure 4 illustrates the correlation between significant engagement behaviours and final scores across clusters in the scatter plot. Clusters 0 and 2 (moderate and high involvement) exhibit a predominantly favourable trend, suggesting that continuous engagement correlates with improved academic performance. Conversely, Cluster 1 displays a more chaotic distribution, wherein certain low-engagement pupils attain unexpectedly high grades.

This mismatch indicates that certain learners may depend significantly on summative examinations while disregarding routine learning activities. This behaviour indicates possible discrepancies between assessment design and learning objectives, highlighting the necessity for instructional practices that promote both formative and summative engagement.

These findings substantiate the educational significance of engagement analytics in identifying at-risk students, tailoring interventions, and enhancing the overall structure of blended learning environments.

Figure 4: Scattergram of Selected Indicators and Scores.



Note: “Cluster 0” represents learners with moderate engagement, “Cluster 1” represents those with low engagement, and “Cluster 2” represents those with high engagement.

4.3. Machine Learning-Based Blended Learning Engagement Assessment Model Construction

4.3.1. Selection of Training and testing Datasets

This research utilises student engagement data gathered across three semesters of a blended learning course. The limited number of students per semester results in a modest dataset, potentially impacting the stability of segmentation-based algorithms like Decision Trees and Random Forests. To address this restriction, 10-fold cross-validation is employed to assess model performance in batch learning. The dataset is randomly divided into ten subgroups; in each iteration, one subset functions as the testing set while the other nine constitute the training set. Performance metrics—accuracy, Macro Precision, Macro Recall, and Macro F1-score—are averaged over the ten folds to yield a dependable evaluation of the model's generalisation capability.

The study employs a chronological division for model evaluation: data from the initial two semesters is utilised for model training, whereas the third semester functions as the independent testing set. Hyperparameter tuning is performed to enhance the efficacy of the Decision Tree and Random Forest algorithms.

4.3.2. Model Construction and Evaluation

Three categorisation models are created using the open-source Scikit-learn toolkit and Python: Decision Tree, Support Vector Machine (SVM), and Random Forest. Two model training methodologies are evaluated to determine their efficacy in predicting engagement.

Model Based on Batch Learning: Scheme 1.

This method consolidates data from the initial two semesters into a unified training set. The data from the third semester is allocated for testing purposes. Table 3 encapsulates the performance indicators for all three models inside this framework.

Table 3: The Accuracy, Macro-P, Macro-R and Macro-F1 scores of the Three Algorithms.

Algorithm	Accuracy	Macro P	Macro R	Macro F1
Decision Tree	0.5918	0.4775	0.5224	0.6050
SVM	0.6939	0.5806	0.6594	0.7126
Random Forest	0.7347	0.5839	0.5900	0.7290

Random Forest performs best in terms of both classification accuracy and balance across precision, recall, and F1-score. Batch learning allows the model to capture general patterns from larger aggregated datasets. However, it may suffer from reduced adaptability when applied to new or evolving data, particularly in educational settings where teaching strategies and learner behaviors change over time.

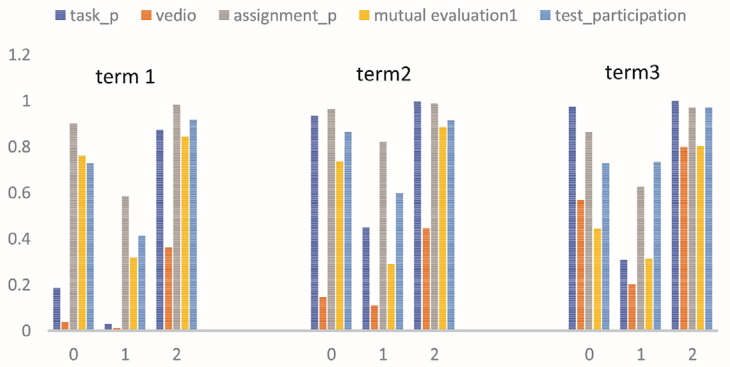
Longitudinal analysis reveals a progressive improvement in learner engagement across semesters, notably in task completion, discussion participation, peer evaluation, and video engagement. This trend reflects both instructional adjustments by teachers and evolving learner behaviors. Figure 5 visualizes the shifts in indicator values by engagement level across semesters.

Scheme 2: Incremental Learning-Based Model.

To address the limitations of batch learning, Scheme 2 employs incremental learning, specifically using the incremental Random Forest algorithm. This approach updates the model iteratively as new data becomes available, enabling better adaptability to cohort changes and real-time application scenarios.

Foro de Educación, v. 23, n. 2, julio-diciembre / july-december 2025, pp. 111-131.
e-ISSN: 1698-7802

Figure 5: Comparison of Engagement Indicators Across Three Semesters.



Note: Number 0 represents the learner with medium engagement, number 1 represents the learner with low engagement, and number 2 represents the learner with high engagement.

In this scheme, training still uses data from the first two semesters, but the datasets are treated separately rather than merged. The third semester remains the testing set. To evaluate model robustness and generalizability, 10-fold cross-validation is again applied. The model is tested under varying data conditions, including different cohort sizes and class imbalances.

Table 4 compares the classification performance of models using both batch and incremental learning, across two training datasets (Semester 1 and Semester 2).

Table 4: Classification Performance of batch vs. Incremental Learning.

Dataset	Model	Accuracy	Macro P	Macro R	Macro F1
Dataset 1	Decision Tree	0.7347	0.7998	0.6641	0.7420
	SVM	0.7143	0.6084	0.7150	0.7318
	Random Forest (batch)	0.7959	0.8140	0.5707	0.7604
	Random Forest (incremental)	0.7755	0.6862	0.5657	0.7429
Dataset 2	Decision Tree	0.4898	0.5479	0.5698	0.5223
	SVM	0.6939	0.5727	0.5691	0.6893
	Random Forest (batch)	0.6939	0.5983	0.7046	0.7152
	Random Forest (incremental)	0.8163	0.7454	0.6768	0.7933

Results show that when trained on Dataset 1, batch Random Forest achieves the highest accuracy. However, when trained on Dataset 2, incremental Random Forest outperforms all others. This demonstrates that incremental learning is more adaptable to evolving learner behaviors and is better suited for dynamic educational contexts.

5. Discussion

This study introduced a machine learning methodology for evaluating student involvement in a dual-subject blended learning context, specifically applied in the Database Principles and Applications course. The study formulated 11 essential indicators that encapsulate the behavioural, cognitive, and social interaction elements of engagement by amalgamating behavioural traces from both online and offline activities, as delineated in Fredricks et al.'s (2004) tripartite model. These markers were integrated into a dynamic evaluation system, facilitating ongoing, data-driven monitoring of student engagement during various phases of education.

Three notable conclusions arose from the empirical analysis:

The K-Means clustering algorithm facilitated the precise categorisation of students into unique interaction profiles. These profiles disclosed significant variations in student participation during the course, providing teachers with detailed insights into learner behaviour. This discovery corresponds with other studies on unsupervised learning methodologies in educational contexts, wherein clustering has demonstrated the ability to reveal hidden engagement patterns that may be neglected by conventional evaluations (Hung et al., 2020; Jovanović et al., 2017). The categorisation of engagement archetypes—namely consistently active, selectively engaged, and minimally involved learners—establishes a pragmatic basis for tailored instructional strategies. Furthermore, these typologies enable educators to perceive involvement not as a dichotomous condition but as a continuum, so facilitating opportunities for formative feedback and innovative educational design.

Secondly, longitudinal examination over three semesters indicated that engagement levels increased when instructors applied iterative pedagogical enhancements. The enhancements encompassed modifications to task complexity, reorganisation of group collaboration, and improved integration of pre-class resources with in-class activities. The findings corroborate prior evidence highlighting that pedagogical adaptation is crucial for maintaining engagement in blended environments (Boelens et al., 2017; Islam et al., 2022). Engagement improvements were more pronounced in groups where instructors actively addressed previous feedback and adjusted their teaching methods accordingly. This highlights the interdependent relationship between instructional design and student engagement—a crucial factor in the development of blended courses. The iterative method of course delivery embodies the principles of action research and practitioner-led inquiry, consistent with the reflective teaching philosophy promoted in critical pedagogical traditions.

The comparison between batch and incremental machine learning models revealed that the Random Forest incremental algorithm exhibited enhanced performance in predicting continued involvement. In contrast to conventional batch models that necessitate full retraining with each new dataset, the incremental approach enables the system to update predictions in near real-time while maintaining model stability and accuracy. This capacity is particularly crucial in educational settings, since engagement is dynamic and can change swiftly due to individual, academic, or contextual influences. The findings correspond with extensive research in educational data mining, which promotes adaptive learning analytics systems that facilitate personalised and responsive instruction (Aldowah et al., 2019; Almusaed et al., 2023). Furthermore, the incorporation of real-time predictive technologies facilitates anticipatory teaching interventions, enabling educators to proactively mitigate disengagement patterns before they become entrenched, so fostering fairness and inclusion in instructional responses.

Notwithstanding these encouraging results, the study possesses specific limitations. A significant disadvantage is the absence of emotional engagement, partly attributable to the dependence on LMS logs and activity records that predominantly document observable behaviours. Considering the increasing focus on affective factors in education, subsequent research should investigate the amalgamation of multimodal data sources—such as facial expression analysis, speech tone recognition, and sentiment analysis from forum posts—to develop a more comprehensive engagement profile (Tang et al., 2025; Wong et al., 2024). This is especially relevant in culturally diverse educational settings, where students' emotional signals and expression methods differ, and where algorithmic biases may compromise detection accuracy if models are not sufficiently localised.

The present study was confined to a singular course context and concentrated on a specific machine learning technique. To improve the generalisability of the results, subsequent research should assess a broader range of incremental algorithms and apply the framework across diverse fields with differing educational needs and student demographics. Collaborations between institutions may be particularly beneficial in developing scalable engagement analytics ecosystems that support diverse curricular and technological frameworks. Moreover, incorporating qualitative inquiry—such as student interviews or teacher reflections—could correlate the engagement profiles with lived experiences and yield better substantiated interpretations of the data.

Ultimately, although the clustering outcomes yielded valuable engagement typologies, the interpretability of machine learning models continues to pose a barrier. Enhancing model transparency, particularly through explainable AI techniques, is anticipated to boost usability and adoption among educational practitioners who are more inclined to employ tools they comprehend. Addressing this issue may reconcile technology innovation with pedagogical practice, facilitating more informed decision-making in actual classroom environments. In addition to technical clarity, this raises ethical considerations: if algorithmic recommendations are to impact instructional decisions, educators must be prepared to scrutinise and evaluate these outputs within a comprehensive framework of professional judgement and educational principles.

This paradigm underscores the wider relevance of machine learning in blended learning research beyond the specific setting of this study. This approach illustrates the transformation of interaction data into actionable insights, highlighting the possibility for real-time, adaptive learning environments that transcend computer science classes. This scalability indicates that the framework may enhance research and teaching methodologies in other fields where blended learning is progressively embraced. This study not only enhances methodological innovation in educational data mining but also establishes a basis for future interdisciplinary investigation of adaptive teaching tactics.

6. Conclusion

This study introduces a systematic and scalable framework for evaluating student engagement in blended learning environments, addressing the rising demand for data-driven teaching practices in higher education. The suggested approach, grounded on the tripartite framework of engagement (Fredricks et al., 2004) and utilising machine learning's predictive capabilities, provides significant theoretical and practical contributions to the expanding domain of learning analytics.

The study emphasises three fundamental contributions:

The application of clustering algorithms facilitates the identification of varied engagement profiles, offering instructors a refined comprehension of student participation.

The implementation of incremental learning models, especially those based on Random Forest algorithms, enables ongoing monitoring and adaptive assistance without necessitating significant reconfiguration.

The approach illustrates how longitudinal data collection can enhance sustainable instructional improvement, in accordance with demands for evidence-based pedagogical innovation in blended and hybrid models (Boelens et al., 2017; Saichaie, 2020).

The framework extends beyond its specific use in the Database Principles and Applications course, possessing wider significance for schools aiming to integrate learning analytics into standard pedagogical practices. The modular design facilitates adaption across many disciplines, teaching methodologies, and technology frameworks. The global expansion of blended learning, propelled by pedagogical objectives and digital transformation, necessitates the urgent development of responsive, student-centered interaction platforms. Systems that accommodate varied learner requirements, particularly those of marginalised or non-traditional students, align with overarching demands for inclusive and justice-focused educational innovation (Adeleye et al., 2024).

Furthermore, by allowing educators to monitor engagement longitudinally and modify techniques as needed, the technology facilitates a transition from reactive to proactive teaching methodologies. This corresponds with modern concepts of personalised learning, wherein education is perpetually modified to address learners' changing requirements (Almusaed et al., 2023; Islam et al., 2022). This technique enhances teacher agency, allowing educators to acquire practical insights rather than merely serving as implementers of algorithmic outputs. They instead assume the role of co-interpreters of learning data, participating in pedagogical decision-making that is informed by, although not dictated by, technological inputs.

The report indicates two possible avenues for future investigation. Integrating emotional and affective data into the engagement model would enhance the existing framework and facilitate a more comprehensive understanding of learning processes. Secondly, implementing the model in transdisciplinary and cross-institutional contexts could evaluate its scalability and facilitate its incorporation into institutional decision-making. Subsequent enquiries may also examine ethical considerations, including data privacy, student consent, and the equilibrium between surveillance and support in analytics-based educational environments.

In addition to formal ethical approval and data anonymisation, the study also addressed vulnerabilities related to algorithmic surveillance. Although the collecting of engagement data provides educational benefits, it poses issues with student privacy and the potential for inadvertent surveillance consequences. To alleviate these dangers, the system was engineered to reduce personally identifiable information and to guarantee that data utilisation remains only educational. Another significant factor is the interpretability of machine learning models. The readiness of educators to implement such systems is contingent upon both predicted accuracy and the transparency of model conclusions. Consequently, interpretable outputs—such as feature importance rankings and visualisations of decision processes—were incorporated into the analytical workflow, allowing educators to make educated assessments of student involvement while maintaining confidence in the system.

This study enhances intelligent engagement assessment by illustrating how machine learning may facilitate more responsive, data-driven, and inclusive blended learning experiences. As digital technologies transform education, these methods provide a means to create more egalitarian, personalised, and effective learning environments that adhere to pedagogical principles and social responsibility.

6.1. Acknowledgements

The authors confirm that this manuscript is original. An AI-based tool was used only for language editing (grammar/clarity). All ideas, data, analyses, and conclusions are entirely the authors' own.

6.2. Funding Statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

6.3. Data Availability Statement

Data sets generated during the current study are available from the corresponding author on reasonable request.

6.4. Conflicts of Interest

The authors declare that they have no competing interest.

6.5. Author Contributions

Conceptualization, Methodology, Algorithm Implementation, Formal Analysis, Investigation, Data Curation, Writing – Original Draft, Visualization, and Project Administration: Zhenfeng Jiang; Validation and Writing – Review & Editing: Aidah Abdul Karim and Fariza Khalid; Supervision: Aidah Abdul Karim.

References

Adeleye, O. O., Eden, C. A. & Adeniyi, I. S. (2024). Innovative Teaching Methodologies in the Era of Artificial Intelligence: A Review of Inclusive Educational Practices. *World Journal of Advanced Engineering Technology and Sciences, 11*(2), pp. 069-079. doi: https://doi.org/10.30574/wjaets.2024.11.2.0091

Aldowah, H., Al-Samarraie, H. & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics, 37*, pp. 13-49. doi: https://doi.org/10.1016/j.tele.2019.01.007

Almusaed, A., Almssad, A., Yitmen, I. & Homod, R. Z. (2023). Enhancing Student Engagement: Harnessing "AIED"'s Power in Hybrid Education—A Review Analysis. *Education Sciences, 13*(7), pp. 632. doi: https://doi.org/10.3390/educsci13070632

Boelens, R., De Wever, B. & Voet, M. (2017). Four key challenges to the design of blended learning: A systematic literature review. *Educational Research Review, 22*, pp. 1-18. doi: https://doi.org/10.1016/j.edurev.2017.06.001

Fredricks, J. A., Blumenfeld, P. C. & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research, 74*(1), pp. 59-109. doi: https://doi.org/10.3102/00346543074001059

García-Peñalvo, F. J., Fidalgo-Blanco, Á. & Sein-Echaluce, M. L. (2018). An adaptive hybrid MOOC model: Disrupting the MOOC concept in higher education. *Telematics and Informatics, 35*(4), pp. 1018-1030. doi: https://doi.org/10.1016/j.tele.2017.09.012

Huang, M., Kuang, F. & Ling, Y. (2022). EFL learners' engagement in different activities of blended learning environment. *Asian-Pacific Journal of Second and Foreign Language Education, 7*(1), pp. 9. doi: https://doi.org/10.1186/s40862-022-00136-7

Hung, H.-C., Liu, I.-F., Liang, C.-T. & Su, Y.-S. (2020). Applying Educational Data Mining to Explore Students' Learning Patterns in the Flipped Learning Approach for Coding Education. *Symmetry, 12*(2), pp. 213. doi: https://doi.org/10.3390/sym12020213

Islam, M. K., Sarker, M. F. H. & Islam, M. S. (2022). Promoting Student-Centred Blended Learning in Higher Education: A Model. *E-Learning and Digital Media, 19*(1), pp. 36-54. doi: https://doi.org/10.1177/20427530211027721

Jovanović, J., Gašević, D., Dawson, S., Pardo, A. & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education, 33*, pp. 74-85. doi: https://doi.org/10.1016/j.iheduc.2017.02.001

Liu, M., Zhao, G., Zhong, Z., Ma, J. & Wang, W. (2024). Theoretical Foundations for Blended Learning. In M. Li, X. Han, & J. Cheng (Eds.), *Handbook of Educational Reform Through Blended Learning* (pp. 1-44). Springer Nature Singapore. doi: https://doi.org/10.1007/978-981-99-6269-3_1

Manwaring, K. C., Larsen, R., Graham, C. R., Henrie, C. R. & Halverson, L. R. (2017). Investigating student engagement in blended learning settings using experience sampling and structural equation modeling. *The Internet and Higher Education, 35*, pp. 21-33. doi: https://doi.org/10.1016/j.iheduc.2017.06.002

McGee, P. & Reis, A. (2012). Blended Course Design: A Synthesis of Best Practices. *Journal of Asynchronous Learning Networks, 16*(4), pp. 7-22. doi: https://doi.org/10.24059/olj.v16i4.239

Nguyen, V. A. (2017). The Impact of Online Learning Activities on Student Learning Outcome in Blended Learning Course. *Journal of Information & Knowledge Management, 16*(04), pp. 1750040. doi: https://doi.org/10.1142/S021964921750040X

Prameswоro, I. S., Fathurrochman, I., Sembing, D., Bangkara, B. M. A. S. A. & Sudrajat, D. (2023). Relevance between Blended Learning and Students' Independent Learning Curriculum: An Overview of Digital Age Education, Student and Teacher Engagement, Technological Resources. *Jurnal Kependidikan : Jurnal Hasil Penelitian dan Kajian Kepustakaan di Bidang Pendidikan, Pengajaran, dan Pembelajaran, 9*(3), pp. 858-869. doi: https://doi.org/10.33394/jk.v9i3.8320

Saichaie, K. (2020). Blended, Flipped, and Hybrid Learning: Definitions, Developments, and Directions. *New Directions for Teaching and Learning, 2020*(164), pp. 95-104. doi: https://doi.org/10.1002/tl.20428

Salas-Pilco, S. Z., Yang, Y. & Zhang, Z. (2022). Student engagement in online learning in Latin American higher education during the COVID-19 pandemic: A systematic review. *British Journal of Educational Technology, 53*(3), pp. 593-619. doi: https://doi.org/10.1111/bjet.13190

Sareen, S. & Mandal, S. (2024). Challenges of blended learning in higher education across global north-south: A systematic and integrative literature review. *Social Sciences & Humanities Open, 10*, pp. 101011. doi: https://doi.org/10.1016/j.ssaho.2024.101011

Sinha, S., Rogat, T. K., Adams-Wiggins, K. R. & Hmelo-Silver, C. E. (2015). Collaborative group engagement in a computer-supported inquiry learning environment. *International Journal of Computer-Supported Collaborative Learning, 10*(3), pp. 273-307. doi: https://doi.org/10.1007/s11412-015-9218-y

Tang, X., Gong, Y., Xiao, Y., Xiong, J. & Bao, L. (2025). Facial Expression Recognition for Probing Students' Emotional Engagement in Science Learning. *Journal of Science Education and Technology, 34*(1), pp. 13-30. doi: https://doi.org/10.1007/s10956-024-10143-7

Wong, Z. Y., Liem, G. A. D., Chan, M. & Datu, J. A. D. (2024). Student Engagement and Its Association with Academic Achievement and Subjective Well-Being: A Systematic Review and Meta-Analysis. *Journal of Educational Psychology, 116*(1), pp. 48-75. doi: https://doi.org/10.1037/edu0000833

Yılmaz, F. G. K. & Yılmaz, R. (2023). Exploring the role of sociability, sense of community and course satisfaction on students' engagement in flipped classroom supported by facebook groups. *Journal of Computers in Education, 10*(1), pp. 135-162. doi: https://doi.org/10.1007/s40692-022-00226-y

Yu, Y., Che Tak, K. B., Bailey, R. P., Samsudin, N. & Ren, C. (2025). The Effects of Blended Learning on Learning Engagement in Physical Education Among University Students in China: The Mediating Role of Attitudes. *Sustainability, 17*(2), pp. 378. doi: https://doi.org/10.3390/su17020378

Zhoc, K. C. H., Webster, B. J., King, R. B., Li, J. C. H. & Chung, T. S. H. (2019). Higher Education Student Engagement Scale (HESES): Development and Psychometric Evidence. *Research in Higher Education, 60*(2), pp. 219-244. doi: https://doi.org/10.1007/s11162-018-9510-6