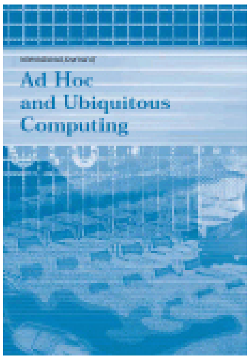
INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research



International Journal of Ad Hoc and Ubiquitous Computing

ISSN online: 1743-8233 - ISSN print: 1743-8225
https://www.inderscience.com/ijahuc

Learning boosting: enhancing predictive modelling in blended learning environments

Zhihong Xu, Chin-Hwa Kuo, Chih-Yung Chang, Jinjun Liu, Chunyan Yu

DOI: 10.1504/IJAHUC.2025.10073446

Article History:

Received:	08 May 2025
Last revised:	24 July 2025
Accepted:	31 July 2025
Published online:	10 March 2026

Learning boosting: enhancing predictive modelling in blended learning environments

Zhihong Xu

School of Computer and Information Engineering,
Chuzhou University,
Anhui, China
Email: xzh@chzu.edu.cn

Chin-Hwa Kuo* and Chih-Yung Chang

Department of Computer Science and Information Engineering,
Tamkang University,
New Taipei 25137, Taiwan
Email: chinhwakuo@mail.tku.edu.tw
Email: cychang@mail.tku.edu.tw
*Corresponding author

Jinjun Liu and Chunyan Yu

School of Computer and Information Engineering,
Chuzhou University,
Anhui, China
Email: jinjunl@chzu.edu.cn
Email: yuchy@chzu.edu.cn

Abstract: Accurate prediction of student learning outcomes is critical for early intervention and instructional decision-making in blended learning environments. This study proposes learning boosting, a structure-enhanced predictive framework integrating community-aware Louvain clustering with a gradient boosting classification. Student activity graphs are clustered to detect latent behavioural communities, and the resulting structural labels are embedded as features for final prediction. Experiments on real-world data from a blended learning course with 102 students evaluate the method under multiple classification granularities, data modalities, and clustering strategies. Results show that learning boosting consistently outperforms 11 baseline models, achieving an F1-score of 0.892, AUC of 0.883, and recall of 0.903 in the three-class task. Ablation studies confirm the complementary benefits of structural feature extraction and clustering. The findings demonstrate that combining graph-based structural modelling with boosting classifiers offers a robust and interpretable approach to learning analytics, especially in sparse and multimodal conditions.

Keywords: learning boosting; learning analytics; blended learning; student performance prediction; gradient boosting.

Reference to this paper should be made as follows: Xu, Z., Kuo, C-H., Chang, C-Y., Liu, J. and Yu, C. (2026) ‘Learning boosting: enhancing predictive modelling in blended learning environments’, *Int. J. Ad Hoc and Ubiquitous Computing*, Vol. 51, No. 2, pp.116–126.

Biographical notes: Zhihong Xu received her MS degree from Anhui University, Hefei, China, in 2008. She is currently an Associate Professor with the School of Computer and Information Engineering, Chuzhou University, Chuzhou, Anhui, China. Her research interests include data analysis and prediction, machine learning, and data visualisation.

Chin-Hwa Kuo received his BS in Mechanical Engineering from Chung-Yuan University, Taoyuan City, Taiwan in 1980, MS in Electrical Engineering from Marquette University, Milwaukee, WI, USA in 1989, and PhD in Electrical Engineering from the University of Notre Dame, South Bend, IN, USA in 1994. He is currently a Professor with the Department of Computer Science and Information Engineering, Tamkang University, New Taipei, Taiwan. His research interests include multimedia processing and applications, learning analytics, and large-scale deployment of information and communications technology in education.

Chih-Yung Chang received his PhD in Computer Science and Information Engineering from the National Central University, Taiwan in 1995. He is currently a Distinguished Professor with the Department of Computer Science and Information Engineering, Tamkang University, New Taipei, Taiwan. His current research interests include artificial intelligence, internet of things, wireless sensor networks, and smart healthcare. He has served as an Associate Guest Editor for several SCI-indexed journals and has published extensively in international journals and conferences in these fields.

Jinjun Liu received his PhD in Computer Application Technology from the School of Computer and Information, Hefei University of Technology, Hefei, China in 2019. He is currently a Professor with the School of Computer and Information Engineering, Chuzhou University, Chuzhou, China. His research interests include gerontechnology, internet of things, and eHealth.

Chunyan Yu received her MS degree from the School of Computer and Information, Hefei University of Technology, Hefei, China in 2008. She is currently a Professor with the School of Computer and Information Engineering, Chuzhou University, Chuzhou, Anhui, China. Her research interests include big data, smart campus, and machine learning.

1 Introduction

With the deep integration of information technology, educational paradigms are shifting from traditional classrooms to blended learning models. By combining self-directed online learning with face-to-face classroom interaction, blended learning offers students personalised learning pathways (Mulenga and Shilongo, 2025). In recent years, increasing attention has been given to predicting student performance, particularly in blended environments where leveraging both online and offline learning data has become a critical research focus (Ragni et al., 2024; Nouri et al., 2019; Anthony et al., 2023). Many existing approaches to student performance prediction are limited to a single learning modality, relying either on online behavioural data or exclusively on classroom performance (Zhang et al., 2023; Katarya, 2023). However, these methods frequently overlook the complementary nature of online and offline learning, resulting in limited prediction accuracy.

To address this issue, recent studies have begun to integrate both online and offline data sources to improve prediction performance (Azizah et al., 2024). Nevertheless, commonly used clustering methods such as K-means and DBSCAN still exhibit limitations when processing heterogeneous data. In response, this study proposes a hybrid analytical framework that integrates hierarchical clustering with ensemble learning techniques. The main contributions of this research are as follows:

- *An innovative hybrid data analysis framework:* this paper introduces, for the first time, a learning boosting-based model that combines community detection with gradient boosting. By clustering student learning behaviours and incorporating the resulting structure into a classification model, the framework addresses the limitations of single-mode data analysis.
- *Optimisation through multimodal data fusion:* the proposed model effectively integrates online and offline data, capitalising on their complementary strengths to enhance prediction accuracy. Compared to traditional single-source models, this approach offers a more comprehensive and efficient prediction strategy.
- *Extensive experimental validation and comparative analysis:* through comprehensive experimentation, this study validates the effectiveness of the proposed method. The Louvain clustering combined with a gradient boosting classifier is compared against other machine learning models, demonstrating superior performance in terms of accuracy and other evaluation metrics

The remainder of the paper is organised as follows: Section 2 reviews related work and analyses existing student performance prediction approaches in blended learning settings, with a focus on the use and limitations of online and offline data. Section 3 presents the proposed method based on Louvain hierarchical clustering and gradient boosting. Section 4 describes the experimental setup, dataset, procedure, and evaluation metrics, followed by comparative results. Section 5 offers an in-depth analysis of findings, discussing advantages, limitations, and comparisons. Finally, Section 6 concludes the paper and outlines directions for future research.

2 Related work

Predicting student performance has emerged as a critical area in educational data mining. Researchers have explored a variety of methods based on online, offline, and blended learning data, leveraging diverse algorithms to improve prediction accuracy.

2.1 Research on online learning data

With the advancement of online learning platforms, online behavioural data such as access logs, assignment submissions, and forum activity have become key predictors of academic outcomes. Machine learning techniques such as support vector machines (SVM), decision trees (DT), random forests (RF) and logistic regression (LR) have been widely used to process such data. For instance, Alshabandar et al. (2020) and Luo et al. (2024) utilised machine learning models on online behavioural datasets to achieve promising prediction results. However, online data are often sparse and feature selection can be challenging. Furthermore, these approaches tend to ignore the influence of offline behaviours, thereby limiting the holistic assessment of student performance (Bernacki et al., 2020; Fahd et al., 2021).

2.2 Research on offline learning data

Offline data, such as attendance records, in-class participation, quizzes, and assignments, can offer a more direct reflection of students' engagement in physical classroom settings. Khan and Ghosh (2021) conducted a systematic review on classroom performance prediction methods, emphasising the importance of behavioural features in offline learning environments. Similarly, Polyzou and Karypis (2019) enhanced prediction accuracy for low-performing students through effective feature extraction. Despite the representativeness of offline data, its collection is resource-intensive and may not fully capture students' learning behaviour in online contexts.

2.3 Research on blended data

In blended learning environments, the integration of online and offline data is gaining traction. Akram et al. (2019) analysed procrastination behaviour based on homework submission data and successfully predicted academic performance through behavioural modelling. Nevertheless, most studies still struggle with effectively fusing multi-source data, and fusion strategies remain underdeveloped, leaving room for optimisation (Liang et al., 2024; Kustitskaya et al., 2022; Chytas et al., 2023; Xu et al., 2021).

2.4 Applications of machine learning algorithms

Machine learning techniques are extensively employed in performance prediction tasks. The gradient boosting algorithm proposed by (Friedman, 2001), along with its variants such as AdaBoost (AdaB), gradient boosting (GB), and CatBoost, has demonstrated superior performance in large-scale data analysis (Kharwar and Thakor, 2023). Studies by (Wang et al., 2022) and (Muhammady et al., 2024) further validated the effectiveness of boosting algorithms in predicting student performance in online learning settings. This study adopts a gradient boosting model, enhanced with Louvain clustering, to improve stability in handling multimodal data.

2.5 Integration of clustering and classification methods

The Louvain algorithm, widely used in social network and educational behaviour analysis, is a prominent method for community detection. Fida et al. (2022) proposed a hybrid ensemble clustering approach that integrates classification with clustering, significantly improving prediction accuracy and validating the utility of structure-enhanced strategies in educational data mining. However, most existing studies rely on single-source data, failing to harness the full potential of blended datasets. Table 1 summarises a comparison between existing methods and the proposed learning boosting model. As illustrated, the integration of clustering with gradient boosting not only enhances feature input by capturing behavioural patterns but also improves performance prediction by modelling student behaviour across learning modalities.

3 Assumptions and problem statement

This section introduces the definitions of notations, assumptions and constraints of this paper. Then, the problem formulation of this study is presented.

3.1 System model and assumptions

This study proposes a method for predicting students' learning assessment outcomes based on online and offline learning activities on the Small Private Online Course (SPOC) platform. It is assumed that learning behaviours consist of online and offline learning activities, and learning activities are a subset of learning behaviours. The specific assumptions are as follows:

Let $S = \{s_1, s_2, \ldots, s_n\}$ be a set of n students, where the learning behaviour of each student s_i is composed of online and offline learning activities. Define the learning behaviour set $B = \{b_1, b_2, \ldots, b_{|B|}\}$, the learning activity set $A = \{a_1, a_2, \ldots, a_{|A|}\}$, and $A \subseteq B$. Learning activities are divided into online and offline activities, defined as follows.

Table 1 Comparisons of the main characteristics of this research with the existing related work

Studies	*Datasets*	*Clustering*	*Boosting*	*Method*
Alshabandar et al. (2020) and Bernacki et al. (2020)	online	No	No	RF, LR, RF, DT, KNN
Fahd et al. (2021)	online	No	Yes	RF, SVM, DT, KNN, AdaB
Wang et al. (2022)	offline	No	Yes	AdaB, XGB, GB
Nouri et al. (2019)	blended	No	No	RF, NB, KNN, DNN, LR
Luo et al. (2024) and Liang et al. (2024)	blended	Yes	No	RF, SVM, kNN, DT
Learning boosting (ours)	blended	Yes	Yes	Louvain + GB

Online activity set $A^{on} = \{x_1^{on}, x_2^{on}, \dots, x_{m_1}^{on}\}$, where x_j^{on} represents the j^{th} type of online learning activity, such as watching online course videos or participating in online forum discussions.

Offline activity set $A^{off} = \{x_1^{off}, x_2^{off}, \dots, x_{m_2}^{off}\}$, where x_k^{off} represents the k^{th} type of offline learning activity, such as class attendance or in-class exercises.

Assume that the learning behaviour record of each student s_i consists of online and offline activities, respectively represented as:

$$B_i^{on} = \left(A_{i,1}^{on}, A_{i,2}^{on}, \dots, A_{i,|B_i^{on}|}^{on}\right)$$

$$B_i^{off} = \left(A_{i,1}^{off}, A_{i,2}^{off}, \dots, A_{i,|B_i^{off}|}^{off}\right)$$

where $|B_i^{on}|$ and $|B_i^{off}|$ represent the number of online and offline learning activities participated in by students, respectively.

The learning assessment outcome R_i is defined as the student's final grade, divided into K levels.

Figure 1 Illustration of the performance assessment structure (see online version for colours)

	x_1^{on}	x_2^{on}	…	x_{m1}^{on}	x_1^{off}	x_2^{off}	…	$x_{m_2}^{off}$	
				B_1					
s_1	$A_{1,1}^{on}$	$A_{1,2}^{on}$	…	A_{1,m_1}^{on}	$A_{1,1}^{off}$	$A_{1,2}^{off}$	…	A_{1,m_2}^{off}	R_1
				B_2					
s_2	$A_{2,1}^{on}$	$A_{2,2}^{on}$	…	A_{2,m_1}^{on}	$A_{2,1}^{off}$	$A_{2,2}^{off}$	…	A_{2,m_2}^{off}	R_2
				…					
…	…								…
				B_n					
s_n	$A_{n,1}^{on}$	$A_{n,2}^{on}$	…	A_{n,m_1}^{on}	$A_{n,1}^{off}$	$A_{n,2}^{off}$	…	A_{n,m_2}^{off}	R_n

3.2 Problem statements

The goal of this study is to predict the learning assessment outcome of students. The prediction model Q takes the students' learning behaviour data B_i^{on} and B_i^{off} as input and predicts the learning assessment outcome. Define the mapping relationship:

$$f: \left(B_i^{on}, B_i^{off}\right) \to R_i, \quad f^Q: \left(B_i^{on}, B_i^{off}\right) \to \hat{R}_i$$

where $\hat{R}_i$ represents the predicted learning assessment outcome, indicating the student's final grade level. R_i is the true value.

For the prediction model Q, its performance is evaluated using the confusion matrix *Conf*, defined as:

$$Conf = Conf_{K \times K}(i, j)$$

where K represents the number of learning assessment levels, and $Conf(i, j)$ represents the number of samples whose true category is i but predicted as j.

Based on the confusion matrix, the following performance metrics are defined:

- True positive (TP):

$$TP_k^Q = Conf(k, k) \tag{1}$$

  Indicates the number of samples correctly predicted as category k by model Q.

- True negative (TN):

$$TN_k^Q = \sum_{1 \le i \le K, 1 \le j \le K, i \ne k, j \ne k} Conf(i, j) \tag{2}$$

  Indicates the number of samples whose true and predicted categories do not belong to k.

- False positive (FP):

$$FP_k^Q = \sum_{1 \le j \le K, i = k, j \ne k} Conf(i, j) \tag{3}$$

  Indicates the number of samples whose true category does not belong to k but are incorrectly predicted as k.

- False negative (FN):

$$FN_k^Q = \sum_{1 \le i \le K, i \ne k, j = k} Conf(i, j) \tag{4}$$

  Indicates the number of samples whose true category belongs to k but are incorrectly predicted as other categories.

Based on the above definitions, the comprehensive evaluation metrics of model Q across all categories are expressed as:

- Total true positive:

$$TP^Q = \sum_{1 \le k \le K} TP_k^Q \tag{5}$$

- Total true negative:

$$TN^Q = \sum_{1 \le k \le K} TN_k^Q \tag{6}$$

- Total false positive:

$$FP^Q = \sum_{1 \le k \le K} FP_k^Q \tag{7}$$

- Total false negative:

$$FN^Q = \sum_{1 \le k \le K} FN_k^Q \tag{8}$$

Based on the global metrics, the precision, recall and F_1-score of the model Q are defined as:

- Precision:

$$Precision^Q = \frac{TP^Q}{TP^Q + FP^Q} \tag{9}$$

- Recall:

$$Recall^Q = \frac{TP^Q}{TP^Q + FN^Q} \tag{10}$$

- Weighted F_1-score:

$$F_1^Q = \alpha F_1^{Q,on} + (1-\alpha) F_1^{Q,off} \tag{11}$$

where α is the weighting coefficient.

3.3 Objective

The objective of this study is to select the optimal prediction model by maximising the model's F_1-score to improve the accuracy of learning assessment outcome predictions. Let Φ be the set of all candidate models, and define the optimal model Q_{best} as:

$$Q_{best} = \arg\max_{Q\in\Phi} F_1^Q$$

where F_1^Q is the F_1-score of the model Q, which balances classification accuracy and the ability to identify minority class samples. The model that maximises F_1^Q is selected as the final best prediction method, Q_{best}.

Additionally, the calculation of the F_1-score considers the weighted combination of online and offline learning activities, ensuring that the model fully utilises all types of learning data to predict students' final grades. When selecting the optimal model, factors such as training time, complexity, and performance in different learning environments are also considered. Therefore, this study aims to explore the impact of various factors on model prediction performance through a series of ablation experiments and ultimately select the best prediction method based on the F_1-score.

4 Proposed mechanism

This section proposes a novel framework for student performance prediction that integrates data preprocessing, feature engineering, clustering technique, and machine learning algorithm. By leveraging blended learning data that combines online and offline activities, the framework aims to accurately predict student academic outcomes. The overall architecture is illustrated in Figure 2.

4.1 Algorithm overview

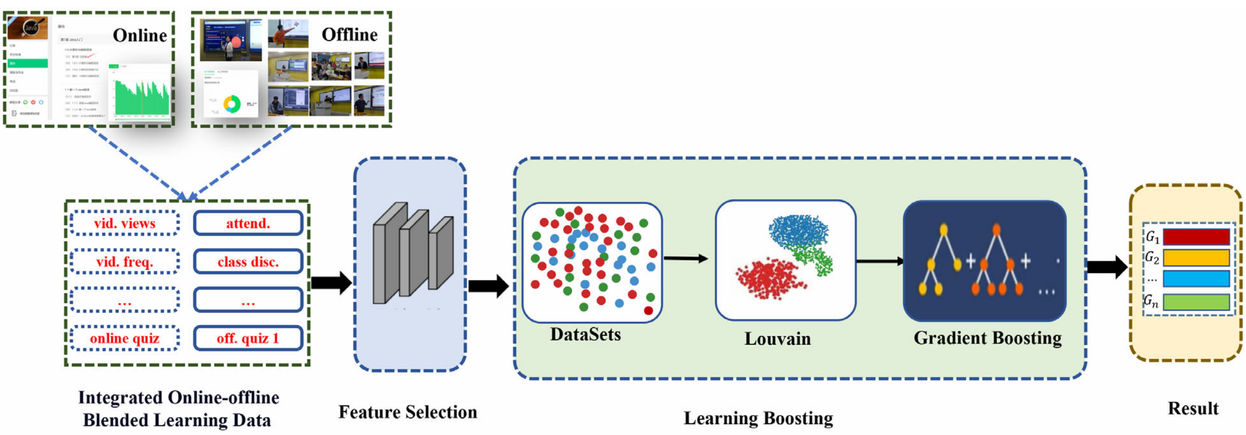
This study proposes a structure-enhanced classification framework, learning boosting, designed to predict student learning outcomes by leveraging both online and offline behavioural data. As shown in Table 2, the process begins with the collection and preprocessing of blended learning data to generate a unified dataset. Then, the information gain ratio is calculated for each feature, and the top-m most informative features are retained to form a reduced dataset D'.

Table 2 The algorithm of learning boosting

	Input: D – integrated online-offline blended learning data
	Output: $F(x)$ – classification model
1	Collect and preprocess online-offline blended learning data, generate dataset D;
2	Calculate information gain ratio for each feature f_i, rank features by information gain ratio, retain top-m most discriminative features, form reduced dataset D';
3	Apply Louvain clustering to D, optimise community structure via modularity and output C clusters: $\{D'_1, \ldots, D'_c\}$;
4	**For** $i = 1$ to C **do**:
5	Train gradient boosting decision tree for each cluster D'_j;
6	Generate cluster-specific model $F(x)$;
7	**end for**

Subsequently, Louvain clustering is applied to uncover latent community structures, optimising modularity and producing C clusters. For each cluster, a dedicated gradient boosting decision tree (GBDT) model is trained to capture group-specific patterns. The final classification system comprises a set of cluster-specific models, collectively forming the predictive function $F(x)$. This modular design enhances both the interpretability and performance of the model under heterogeneous learning behaviour conditions.

Figure 2 System components of the learning boosting (see online version for colours)



4.2 Ranking and selecting features

The primary objective of feature selection is to identify the most informative subset of features from raw data to enhance model performance and improve the interpretability of results. Feature ranking serves as a fundamental step in feature selection, which quantifies the relevance or importance of each feature with respect to the target variable, thereby guiding subsequent feature screening. Unlike direct subset selection methods (e.g., wrapper approaches), ranking-based feature selection belongs to the filter category, offering computational efficiency and independence from downstream machine learning models, making it particularly suitable for high-dimensional datasets such as student learning data.

This paper employs the information gain ratio for feature selection. The gain ratio is computed for each feature according to the following formula:

$$IGR(f_i) = \frac{\textit{Information gain}(f_i)}{\textit{Split information}(f_i)} \tag{13}$$

where information gain measures the reduction in entropy after splitting the dataset using feature f_i reflecting purity improvement. Split information penalises features with excessive branches, ensuring balanced partitioning. Thus, a high gain ratio indicates strong discriminatory power while avoiding overfitting due to excessive splits.

This paper sorts features in descending order of their gain ratio values, selects the top-m features with the highest gain ratio to form the refined dataset D'.

4.3 Clustering learning data

This paper introduces an innovative approach by positioning clustering algorithms prior to classifiers, grouping similar samples into homogeneous clusters while maintaining inter-cluster separability. Dedicated classification models are then trained on each cluster to enhance specificity. The Louvain algorithm, a modularity optimisation-based community detection method, is employed for efficient network data clustering. Unlike traditional methods such as K-means, hierarchical clustering, and DBSCAN, Louvain automatically identifies densely connected groups without requiring a predefined cluster count, proving particularly effective in detecting student groups with similar learning behaviours.

The Louvain algorithm operates through two iterative phases until modularity convergence is achieved. Phase 1 (modularity optimisation) initialises each node (representing a student $s_i \in S$) as an independent community, computes the modularity gain from node movement, and reassigns nodes to maximise the modularity gain until no further improvement is possible. Phase 2 (network aggregation) merges nodes within the same community into super-nodes, reconstructs edge weights, and reapplies phase 1 to the condensed network until modularity stabilises. This two-phase framework ensures an efficient and scalable clustering solution, making it highly suitable for educational dataset analysis.

4.4 Training multiple classifiers

GBDT is a powerful ensemble learning-based predictive model that constructs a strong classifier by iteratively combining multiple weak decision trees. The core principle of this algorithm lies in adopting a gradient descent strategy, where each newly added decision tree fits the prediction residuals of preceding models, with the final output obtained through weighted aggregation of all weak learners' predictions. Compared to conventional decision trees, GBDT employs an additive training paradigm and shrinkage mechanism, enabling it to effectively capture complex nonlinear relationships in student behavioural data while preventing overfitting. Notably, when addressing multi-class classification tasks, GBDT typically utilises a combined approach of softmax function and cross-entropy loss function.

Using the GBDT algorithm, this paper trains a classifier $F_k(x)$ for each grade class ($k \in \{G_1, G_2, G_3\}$) through multiple iterations (resulting in a total of K classifiers). The classifier is trained to minimise the cross-entropy loss function.

$$L = -\sum_{i=1}^{N}\sum_{k=1}^{K} y_{ik}(x_i)\log p_{ik}(x_i) \tag{14}$$

Here, $y_{ik}(x_i)$ is the true label of sample x_i, $p_{ik}(x_i)$ represents the softmax probability.

$$p_{ik}(x_i) = \frac{e^{F_k(x_i)}}{\sum_{j=1}^{K} e^{F_j(x_i)}} \tag{15}$$

The construction process of the k^{th} classifier $F_k(x)$ is as follows: let the number of iterations be M. In the m^{th} iteration $m \in [1, M]$, a decision tree $h_{km}(x)$ is generated for the k^{th} class ($k \in \{G_1, G_2, G_3\}$), where the decision tree fits the negative gradient set $\{r_{ik}^{(m)}(x_i)\}_{i=1}^{N}$ of the current model and progressively corrects prediction errors.

$$r_{ik}^{(m)}(x_i) = -\left.\frac{\partial L}{\partial F_k(x_i)}\right|_{F=F_{m-1}} = y_{ik}(x_i) - p_{ik}^{(m-1)}(x_i) \tag{16}$$

The weighted cumulative formula of the classifier after the m^{th} iteration is shown below, where η is the learning rate.

$$F_{km}(x) = F_{km-1}(x) + \eta h_{km}(x) \tag{17}$$

After the iterations are completed, M decision trees for the k^{th} class are aggregated to form the final predictive classifier $F_k(x) = F_{kM}(x)$. The classifiers for the other classes are built in the same manner, resulting in a total of K classifiers.

For a given sample x_i, the cumulative score $F_k(x_i)$ for the k^{th} class is obtained by traversing all decision trees in its corresponding classifier. Similarly, the scores for the other classes are computed, yielding K scores in total. Each score is then converted into a probability via the softmax function, and the class with the highest probability is selected as the final prediction.

5 Simulation

To assess the effectiveness of the proposed model in predicting student academic performance, we conducted experiments based on real-world data from an undergraduate SPOC platform. The subsequent sections provide details regarding the simulation environment and present the results.

5.1 Data description

The dataset encompasses the online and offline learning behaviours of 102 students. A total of 17 features were extracted and categorised into two main types: online and offline. The online features include indicators such as video viewing frequency, online quiz scores, forum participation, platform-generated composite scores, and programming assignment completion. The offline features consist of attendance records, post-class exercises, in-class participation, lab reports, bonus task completion, two classroom quizzes, and final project scores. These features are depicted on the left side of Figure 3.

To facilitate modelling, students' final grades were divided into three categories, ensuring balanced class distribution, as shown in Table 3.

To explore differences in group learning behaviours, the Louvain algorithm was employed for unsupervised clustering. The analysis revealed that students could be grouped into three distinct communities, each representing a different pattern of learning engagement. To enhance interpretability and reduce feature dimensionality, we employed information gain to select seven representative features, including programming assignments, total blended scores (online and offline), two offline quizzes, final project score, attendance, and bonus tasks. These selected features are illustrated on the right side of Figure 3.

Table 3 Distribution of student grades

Grade range	*Criteria*
G_1	Score < 60
G_2	60 ≤ Score < 84
G_3	Score ≥ 85

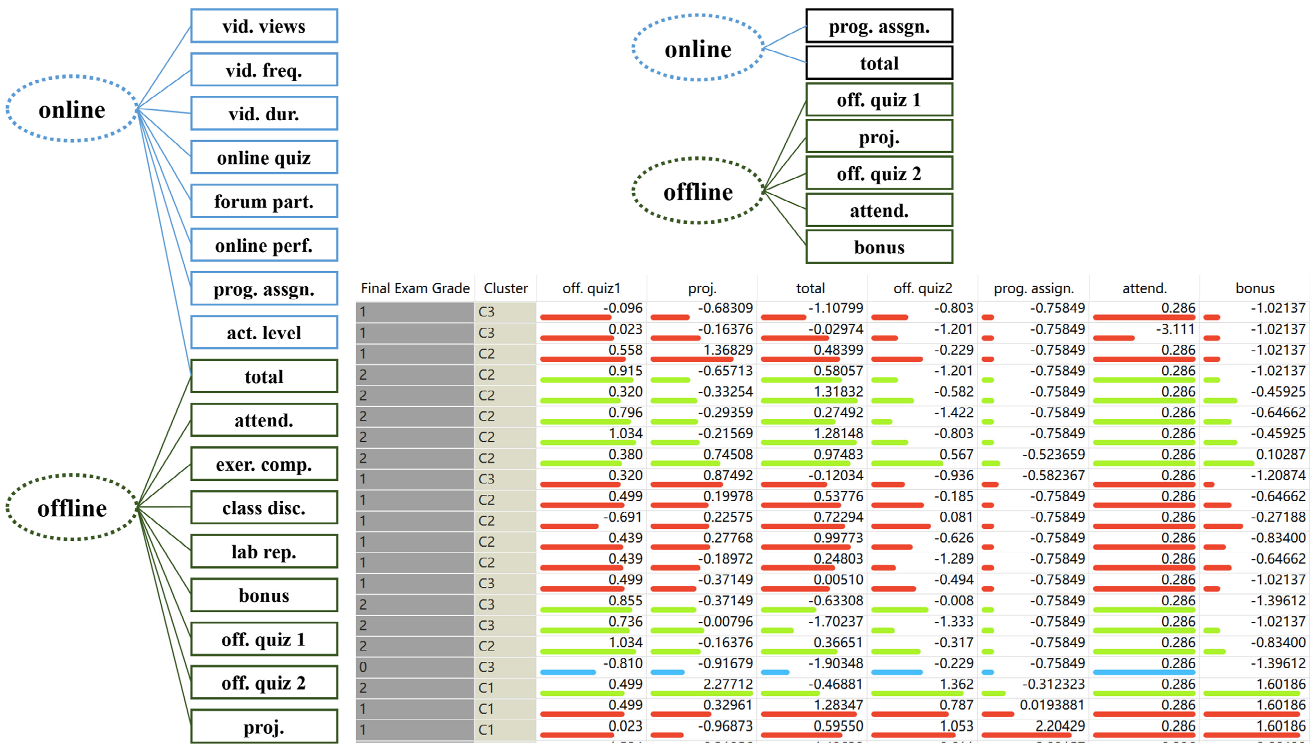
5.2 Experimental results

To evaluate the performance of the learning boosting model under different experimental settings, we conducted 10-fold cross-validation and measured multiple performance metrics, including accuracy, precision, recall, F_1-score, area under the curve (AUC), and Matthews correlation coefficient (MCC).

5.2.1 Classification without clustering mechanism

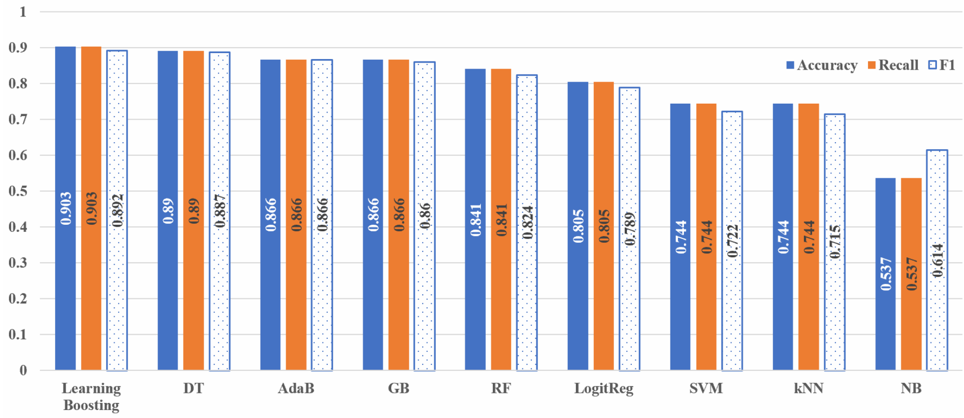
As shown in Figure 4, traditional classifiers without structural enhancement achieved varying levels of performance. Among them, the DT model performed best, reaching an F_1-score of 0.887, followed by AdaBoost and random forest, both exceeding 0.82.

Figure 3 Features and feature-extracted data (see online version for colours)



Final Exam Grade	Cluster	off. quiz1	proj.	total	off. quiz2	prog. assign.	attend.	bonus
1	C3	-0.096	-0.68309	-1.10799	-0.803	-0.75849	0.286	-1.02137
1	C3	0.023	-0.16376	-0.02974	-1.201	-0.75849	-3.111	-1.02137
1	C2	0.558	1.36829	0.48399	-0.229	-0.75849	0.286	-1.02137
2	C2	0.915	-0.65713	0.58057	-1.201	-0.75849	0.286	-1.02137
2	C2	0.320	-0.33254	1.31832	-0.582	-0.75849	0.286	-0.45925
2	C2	0.796	-0.29359	0.27492	-1.422	-0.75849	0.286	-0.64662
2	C2	1.034	-0.21569	1.28148	-0.803	-0.75849	0.286	-0.45925
2	C2	0.380	0.74508	0.97483	0.567	-0.523659	0.286	0.10287
1	C3	0.320	0.87492	-0.12034	-0.936	-0.582367	0.286	-1.20874
1	C2	0.499	0.19978	0.53776	-0.185	-0.75849	0.286	-0.64662
1	C2	-0.691	0.22575	0.72294	0.081	-0.75849	0.286	-0.27188
1	C2	0.439	0.27768	0.99773	-0.626	-0.75849	0.286	-0.83400
1	C2	0.439	-0.18972	0.24803	-1.289	-0.75849	0.286	-0.64662
1	C3	0.499	-0.37149	0.00510	-0.494	-0.75849	0.286	-1.02137
2	C3	0.855	-0.37149	-0.63308	-0.008	-0.75849	0.286	-1.39612
2	C3	0.736	-0.00796	-1.70237	-1.333	-0.75849	0.286	-1.02137
2	C2	1.034	-0.16376	0.36651	-0.317	-0.75849	0.286	-0.83400
0	C3	-0.810	-0.91679	-1.90348	-0.229	-0.75849	0.286	-1.39612
2	C1	0.499	2.27712	-0.46881	1.362	-0.312323	0.286	1.60186
1	C1	0.499	0.32961	1.28347	0.787	0.0193881	0.286	1.60186
1	C1	0.023	-0.96873	0.59550	1.053	2.20429	0.286	1.60186

Figure 4 Scores of CWCM (see online version for colours)



Although DT-based models performed well on datasets of moderate size, the absence of structural information limited their ability to capture community-level behaviour patterns among students, especially when modelling complex feature interactions. In contrast, the learning boosting model achieved a higher F_1-score of 0.892 under the same task, indicating that the incorporation of structural clustering significantly enhances model expressiveness and generalisation.

5.2.2 Clustering then classify mechanism

Table 4 presents the results of combining different clustering methods (Louvain, hierarchical, K-means, DBSCAN) with various classifiers. Among all combinations, Louvain + gradient boosting achieved the highest F_1-score of 0.892 and demonstrated the most stable performance across all metrics.

Louvain clustering effectively identified natural student communities based on graph structure, producing meaningful structural labels. When combined with the robust generalisation capabilities of gradient boosting, this integration led to significant performance improvements. Although K-means + AdaBoost also showed competitive results, it was less effective in modelling fine-grained class boundaries compared to learning boosting.

5.2.3 Clustering-then-neural network mechanism

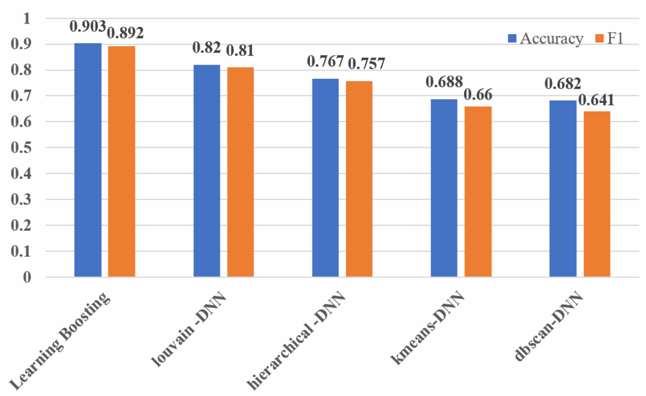
As shown in Figure 5, the Louvain-DNN combination achieved the best performance among all neural configurations, with an F_1-score of 0.810.

However, it still lagged behind learning boosting. While DNNs are capable of capturing complex nonlinear relationships, they tend to be unstable in small-to-moderate data scenarios. Moreover, structural labels derived from Louvain clustering may not be fully leveraged in DNN architectures, explaining their inferior performance relative to tree-based models.

Table 4 Scores of CTCM

Clustering	*Model*	*Accuracy*	*Precision*	*Recall*	F_1
K-means	AdaB	0.889	0.898	0.889	0.892
	kNN	0.715	0.610	0.715	0.653
	LR	0.789	0.683	0.789	0.726
	NB	0.226	0.322	0.226	0.195
	RF	0.789	0.682	0.789	0.727
	SVM	0.782	0.676	0.782	0.719
	DT	0.710	0.665	0.710	0.682
	GB	0.762	0.675	0.762	0.709
DBSCAN	AdaB	0.599	0.611	0.610	0.616
	kNN	0.650	0.610	0.592	0.600
	LR	0.720	0.706	0.690	0.680
	NB	0.679	0.275	0.351	0.679
	RF	0.727	0.718	0.702	0.698
	SVM	0.674	0.706	0.691	0.685
	DT	0.715	0.694	0.694	0.696
	GB	0.667	0.670	0.661	0.656
Hierarchical	AdaB	0.756	0.783	0.756	0.762
	kNN	0.708	0.643	0.708	0.669
	LR	0.805	0.760	0.805	0.778
	NB	0.412	0.749	0.412	0.414
	RF	0.829	0.771	0.829	0.797
	SVM	0.805	0.758	0.805	0.766
	DT	0.817	0.826	0.817	0.816
	GB	0.792	0.787	0.792	0.780
Louvain	AdaB	0.879	0.880	0.879	0.879
	kNN	0.771	0.745	0.771	0.754
	LR	0.856	0.824	0.856	0.839
	NB	0.432	0.684	0.432	0.480
	RF	0.868	0.853	0.868	0.856
	SVM	0.747	0.718	0.747	0.724
	DT	0.879	0.873	0.879	0.874
	Learning boosting (ours)	0.903	0.883	0.903	0.892

Figure 5 Scores of CTNNM (see online version for colours)



5.2.4 *Impact of data modalities on prediction performance*

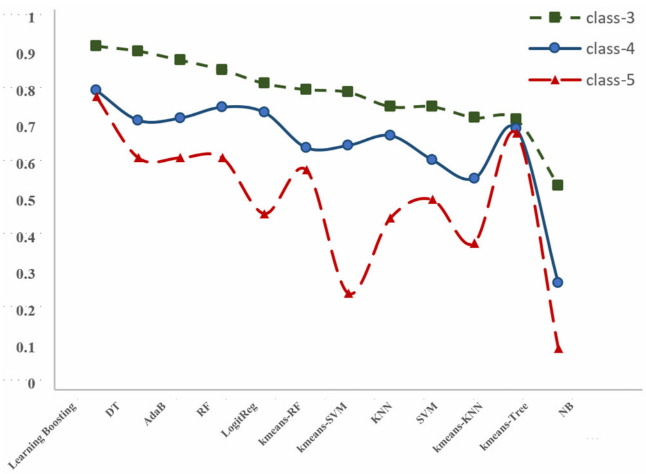
As shown in Table 5, the blended dataset yielded the highest F_1-score of 0.892, followed by offline data at 0.832, while online data alone performed worst at 0.470.

Table 5 Comparison of learning boosting across three data configurations

Datasets	*Accuracy*	*Precision*	*Recall*	F_1
Online	0.479	0.466	0.479	0.470
Offline	0.847	0.821	0.847	0.832
Blended	0.903	0.883	0.903	0.892

Offline data typically contains richer behavioural signals that reflect actual student engagement, whereas online data is often noisier. The integration of both sources enables Louvain clustering to discover more informative structural patterns, thereby enhancing the model's discriminative power.

Figure 6 Comparison of recall rates (see online version for colours)



5.2.5 *Model performance under multi-class classification settings*

We also evaluated model performance under multi-class classification settings, varying the number of grade categories (3, 4 and 5 classes). As illustrated in Figure 6, learning boosting consistently outperformed baseline classifiers including DT, RF, SVM, kNN, LR, NB, AdaBoost and GB (see Table 4) across all configurations.

Even as classification complexity increased with finer grade granularity, learning boosting maintained a high level of performance, with F_1-scores of 0.903, 0.787, and 0.771 for class-3, class-4, and class-5 setups, respectively. This robustness is attributed to the model's ability to adaptively capture structural variations and manage decision boundaries effectively.

5.2.6 *Comparative performance analysis with existing methods*

As depicted in Figure 7, learning boosting consistently outperformed all baseline models across four critical evaluation metrics: accuracy (0.903), recall (0.903), F_1-score (0.892) and Matthews correlation coefficient (MCC = 0.808). These metrics collectively reflect the model's ability to achieve high overall correctness, effective sensitivity to true positives, and balanced performance under class imbalance. This demonstrates its robustness in educational prediction tasks.

In contrast, although decision trees, AdaBoost, and random forest offered competitive precision or AUC in isolated cases, their performance across the four primary metrics remained consistently lower than that of learning boosting. Traditional classifiers such as SVM, logistic regression, and kNN underperformed, likely due to limited capability in capturing nonlinear interactions and structural dependencies.

Additionally, clustering-enhanced models utilising K-means (e.g., K-means + RF, K-means + SVM) failed to demonstrate any structural advantage, highlighting the limitations of distance-based clustering in learning behaviour analysis. By contrast, learning boosting integrates structural community detection with boosting-based classification, enabling more expressive behaviour modelling and superior generalisation in blended learning environments.

5.2.7 *Ablation study*

Table 6 presents the results of an ablation study evaluating the contribution of different components within the proposed learning boosting model. The complete model, consisting of the structural feature extractor, Louvain clustering, and GBDT, achieved the highest F_1-score of 0.892. The removal of any individual module resulted in a noticeable performance decline.

Figure 7 Comparison on key components (see online version for colours)

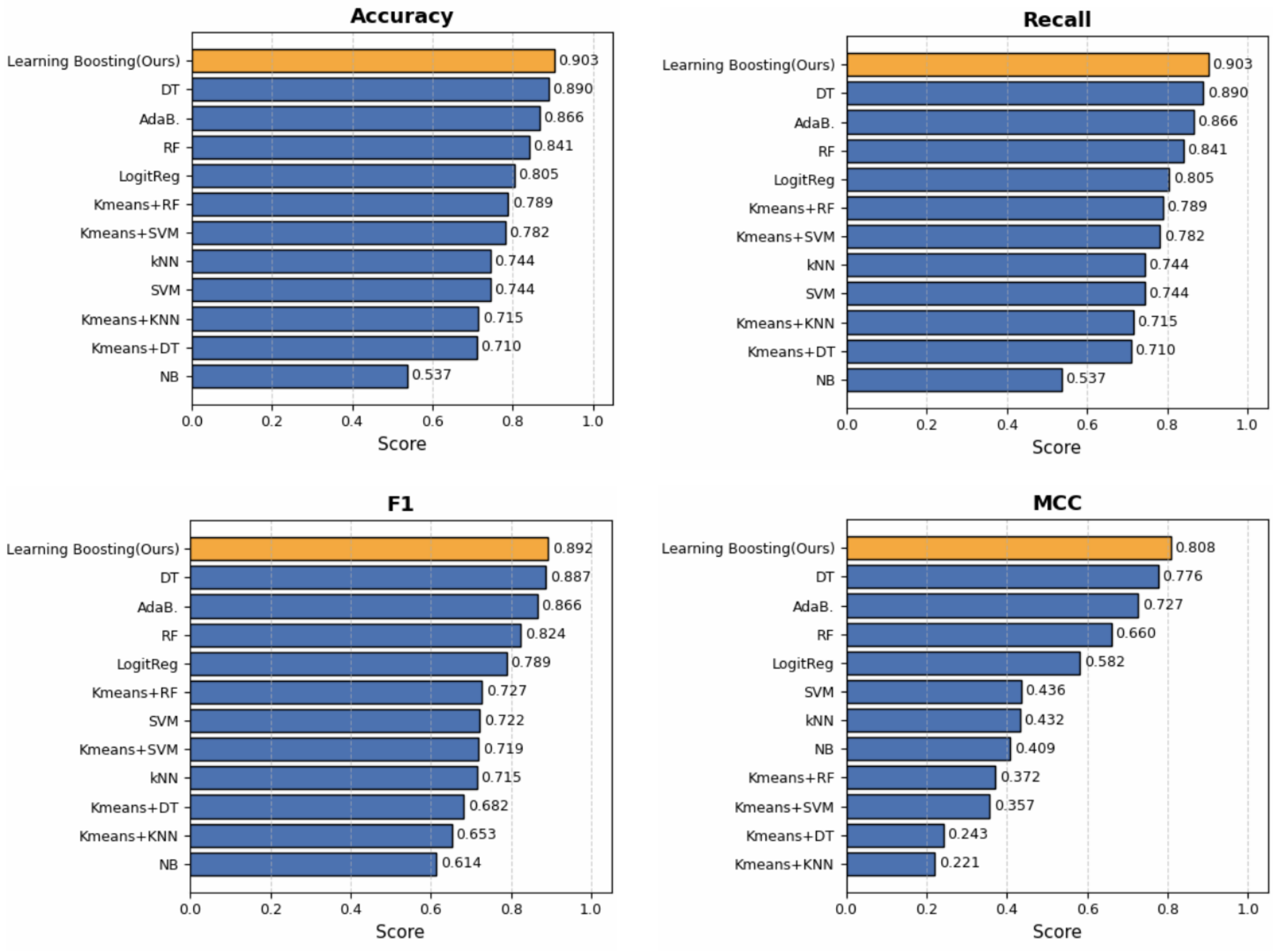


Table 6 Ablation study results

Feature extractor	*Clustering*	*Gradient boosting*	*Accuracy*	*Precision*	*Recall*	F_1
×	×	√	0.759	0.600	0.642	0.740
×	√	√	0.766	0.832	0.823	0.820
√	×	√	0.866	0.855	0.866	0.860
√	√	√	0.903	0.883	0.903	0.892

Among all components, the structural feature extractor was found to have the most significant impact on model performance, followed by the clustering module. The combination of all three modules offers the most comprehensive information representation and the strongest predictive capability. These findings clearly demonstrate the synergistic effect of structural modelling and ensemble learning, highlighting the importance of each module in enhancing the model's effectiveness.

6 Conclusions

This study proposes a structure-enhanced predictive approach, learning boosting, for evaluating student academic performance. The method integrates the community-aware Louvain clustering algorithm with a gradient boosting classifier and is applied to behavioural data modelling in blended learning environments. By extracting latent community structures from student behaviour graphs and embedding them as structural features into the gradient boosting model, the approach effectively combines global behavioural relationships with localised activity representations. Across various classification granularities, data modalities, and comparative baselines, learning boosting consistently exhibits superior predictive performance, outperforming existing methods in key metrics such as F_1-score, AUC, and MCC. The results of the ablation study further confirm that structural feature extraction and clustering modules are critical to the observed performance improvements.

Future work will focus on several directions. First, we aim to construct dynamic behaviour graphs to support temporal modelling. Second, we will explore the incorporation of graph neural networks to enhance the representational power of structural information. Finally, we intend to investigate cross-course and cross-platform transfer and generalisation mechanisms to further improve the model's practicality and deployment scalability.

Declarations

This work was supported in part by the Key Project of Anhui Provincial Research Planning Project No. 2022AH040153, No. 2024xnjys024; Chuzhou University Higher Education Research Project No. 2024jyz024, No. 2024gdjy02.

All authors declare that they have no conflicts of interest.

References

Akram, A., Fu, C., Li, Y., Javed, M.Y., Lin, R., Jiang, Y. and Tang, Y. (2019) 'Predicting students' academic procrastination in blended learning courses using homework submission data', *IEEE Access*, Vol. 22, No. 6, pp.1456–1468.

Alshabandar, R., Hussain, A., Keight, R. and Khan, W. (2020) 'Students performance prediction in online courses using machine learning algorithms', *Proc. of the 2020 Int. Joint Conf. on Networks (IJCNN)*, IEEE, pp.1–7.

Anthony Jr., B., Kamaludin, A. and Romli, A. (2023) 'Predicting academic staff's behaviour intention and actual use of blended learning in higher education: model development and validation', *Technology, Knowledge and Learning*, Vol. 28, No. 3, pp.1223–1269.

Azizah, Z., Ohyama, T., Zhao, X., Ohkawa, Y. and Mitsuishi, T. (2024) 'Predicting at-risk students in the early stage of a blended learning course via machine learning using limited data', *Computers and Education: Artificial Intelligence*, Vol. 7, Article 100261, DOI: 10.1016/j.caeai.2024.100261.

Bernacki, M.L., Chavez, M.M. and Uesbeck, P.M. (2020) 'Predicting achievement and providing support before STEM majors begin to fail', *Computers and Education*, Vol. 158, Article 103999, DOI: 10.1016/j.compedu.2020.103999.

Chytas, K., Tsolakidis, A., Triperina, E. and Skourlas, C. (2023) 'Educational data mining in the academic setting: employing the data produced by blended learning to ameliorate the learning process', *Data Technologies and Applications*, Vol. 57, No. 3, pp.366–384.

Fahd, K., Miah, S.J. and Ahmed, K. (2021) 'Predicting student performance in a blended learning environment using learning management system interaction data', *Applied Computing and Informatics*, online ahead of print, DOI: 10.1108/ACI-06-2021-0150.

Fida, S., Masood, N., Tariq, N. and Qayyum, F. (2022) 'A novel hybrid ensemble clustering technique for student performance prediction', *Journal of Universal Computer Science*, Vol. 28, No. 8, pp.777–779.

Friedman, J.H. (2001) 'Greedy function approximation: a gradient boosting machine', *Annals of Statistics*, Vol. 29, No. 5, pp.1189–1232.

Katarya, R. (2023) 'A systematic review on predicting the performance of students in higher education in offline mode using machine learning techniques', *Wireless Personal Communications*, Vol. 133, No. 3, pp.1643–1674.

Khan, A. and Ghosh, S.K. (2021) 'Student performance analysis and prediction in classroom learning: a review of educational data mining studies', *Education and Information Technologies*, Vol. 26, No. 1, pp.205–240.

Kharwar, A. and Thakor, D. (2023) 'A hybrid approach for feature selection using SFFS and SBFS with extra-tree and classification using XGBoost', *International Journal of Ad Hoc and Ubiquitous Computing*, Vol. 43, No. 4, pp.191–205.

Kustitskaya, T.A., Kytmanov, A.A. and Noskov, M.V. (2022) 'Early student-at-risk detection by current learning performance and learning behavior indicators', *Cybernetics and Information Technologies*, Vol. 22, No. 1, pp.117–133.

Liang, J., Zhao, Z., Ma, S. and Wang, J. (2024) 'Analysing and predicting learning behaviours in computer science: a case study of the blended teaching mode in digital signal processing course', *Journal of Computational Methods in Science and Engineering*, Vol. 24, No. 3, pp.1341–1353.

Luo, Y., Han, X. and Zhang, C. (2024) 'Prediction of learning outcomes with a machine learning algorithm based on online learning behavior data in blended courses', *Asia Pacific Education Review*, Vol. 25, No. 2, pp.267–285.

Muhammady, D.N., Nugraha, H.A.E., Nastiti, V.R.S. and Aditya, C.S.K. (2024) 'Students' final academic score prediction using boosting regression algorithms', *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI)*, Vol. 10, No. 1, pp.154–165.

Mulenga, R. and Shilongo, H. (2025) 'Hybrid and blended learning models: innovations, challenges, and future directions in education', *Acta Pedagogia Asiana*, Vol. 4, No. 1, pp.1–13.

Nouri, J., Saqr, M. and Fors, U. (2019) 'Predicting performance of students in a flipped classroom using machine learning: towards automated data-driven formative feedback', *Proc. of the 10th Int. Conf. on Education, Training and Informatics (ICETI2019)*, Orlando, USA, pp.17–21.

Polyzou, A. and Karypis, G. (2019) 'Feature extraction for next-term prediction of poor student performance', *IEEE Transactions on Learning Technologies,* Vol. 12, No. 2, pp.237–248.

Ragni, A., Ippolito, D. and Masci, C. (2024) 'Assessing the impact of hybrid teaching on students' academic performance via multilevel propensity score-based techniques', *Socio-Economic Planning Sciences*, Vol. 92, Article 101824, DOI: 10.1016/j.seps.2024.101824.

Wang, C., Chang, L. and Liu, T. (2022) 'Predicting student performance in online learning using a highly efficient gradient boosting decision tree', *Proc. of the 2022 Int. Conf. on Intelligent Information Processing*, New York, USA, pp.508–521.

Xu, Z., Yuan, H. and Liu, Q. (2021) 'Student performance prediction based on blended learning', *IEEE Transactions on Education*, Vol. 64, No. 1, pp.66–73.

Zhang, S., Che, S.P., Nan, D. and Kim, J.H. (2023) 'How does online social interaction promote students' continuous learning intentions?', *Frontiers in Psychology*, Vol. 14, Article 1098110, DOI: 10.3389/fpsyg.2023.1098110.