



Learning Analytics-Driven Student Performance Prediction Using Ensemble Machine Learning Models

Elissa Dwi Lestari^{1,*}, Florentina Kurniasari²

¹ Department of Management, Universitas Multimedia Nusantara, Scientia Boulevard Gading, Tangerang 15810, Indonesia

² Magister of Management Technology, Universitas Multimedia Nusantara, Scientia Boulevard Gading, Tangerang 15810, Indonesia

ABSTRACT

The increasing availability of data from digital learning environments has created new opportunities to analyse and predict learning performance using artificial intelligence. Learning analytics enables the extraction of meaningful behavioural patterns from temporal data to support data-driven educational decision-making. This study proposes learning analytics-driven framework for predicting short-term performance outcomes using ensemble machine learning models. The framework is evaluated on a dataset consisting of 1,334 temporal observations, integrating performance indicators, activity intensity measures, volatility metrics, and temporal lag features to capture both instantaneous and sequential learning behaviours. Five machine learning models—Decision Tree, Random Forest, Gradient Boosting, XGBoost, and a Voting Ensemble—are systematically compared using Accuracy, Precision, Recall, and F1-score as evaluation metrics. The experimental results show that the Decision Tree model achieves the highest F1-score of 0.546, while the Voting Ensemble model attains a balanced performance with an accuracy of 0.510 and an F1-score of 0.502, indicating stable classification behaviour across performance improvement and decline classes. Feature importance analysis reveals that short-term performance change indicators and recent activity intensity are the most influential factors in determining subsequent performance outcomes. Although predictive accuracy remains moderate due to the dynamic and volatile nature of temporal performance trajectories, the proposed framework demonstrates the potential of interpretable, learning analytics-driven ensemble models to support early performance tendency detection and adaptive learning interventions in dynamic learning environments.

Keywords Learning Analytics, Ensemble Machine Learning, Student Performance Prediction, Temporal Behaviour Modelling, Artificial Intelligence in Learning

Submitted 5 January 2026
Accepted 9 February 2026
Published 1 March 2026

Corresponding author
Elissa Dwi Lestari,
elissa.lestari@umn.ac.id

Additional Information and
Declarations can be found on
[page 13](#)

DOI: [10.63913/ail.v2i1.5](https://doi.org/10.63913/ail.v2i1.5)

© Copyright
2026 Lestari and Kurniasari

Distributed under
Creative Commons CC-BY 4.0

Introduction

The rapid expansion of digital learning environments has led to an unprecedented growth in educational data generated through online learning platforms, learning management systems, and intelligent tutoring systems. These data provide valuable opportunities to analyse learning behaviours, monitor academic progress, and support data-driven decision-making in education [1]. However, despite the availability of large-scale educational data, many institutions still rely on retrospective evaluation methods that identify learning problems only after performance deterioration has already occurred [2].

Learning analytics has emerged as a prominent research area that focuses on collecting, measuring, and analyzing learner data to understand and optimize learning processes and environments [3]. By leveraging learning analytics,

How to cite this article: E. D. Lestari and F. Kurniasari, "Learning Analytics-Driven Student Performance Prediction Using Ensemble Machine Learning Models," *Artif. Intell. Learn.*, vol. 2, no. 1, pp. 1-16, 2026.

educators and institutions can gain insights into student engagement, learning behaviour, and performance trajectories, enabling more proactive and personalized educational interventions [4]. In recent years, artificial intelligence (AI) and machine learning techniques have been increasingly integrated into learning analytics to enhance predictive capabilities and automate the identification of at-risk learners [5].

Prior studies on student performance prediction have explored a wide range of machine learning approaches, including logistic regression, decision trees, support vector machines, and neural networks [6], [7]. These models have demonstrated promising results in predicting academic outcomes using features such as demographic attributes, assessment scores, and interaction logs from learning management systems [8]. More recent research has shifted toward ensemble learning methods, such as Random Forest and Gradient Boosting, which combine multiple learners to improve robustness and generalization performance [9].

In parallel, temporal modelling has gained increasing attention in learning analytics research. Sequential models incorporating time-dependent features, lag variables, and behavioural transitions have been shown to capture learning dynamics more effectively than static feature representations [10], [11]. Deep learning architectures, including recurrent neural networks and long short-term memory (LSTM) models, have also been applied to model learning sequences and engagement patterns [12]. Despite these advances, many existing studies focus primarily on maximizing predictive accuracy, often at the expense of interpretability and practical applicability in real educational settings [13].

Although ensemble machine learning and temporal learning analytics have shown strong potential, several research gaps remain. First, existing studies frequently emphasize overall prediction accuracy without adequately addressing the inherent uncertainty and volatility of short-term performance outcomes in dynamic learning environments [14]. As a result, model performance is often overestimated, and limited attention is given to balanced classification behaviour across different performance states.

Second, many performance prediction frameworks rely on static or aggregated features, which fail to fully exploit short-term behavioural changes and temporal dependencies that are critical for understanding learning progression [15]. The integration of short-term performance change indicators, volatility measures, and lag-based features remains underexplored, particularly within interpretable ensemble learning frameworks.

Third, while deep learning models have demonstrated strong representational power, their complexity and lack of transparency pose challenges for adoption in educational contexts, where explainability and actionable insights are essential for instructors and decision-makers [16]. There is therefore a need for learning analytics-driven AI approaches that balance predictive capability, interpretability, and practical relevance.

To address these gaps, this study proposes learning analytics-driven framework for predicting short-term performance outcomes using ensemble machine learning models. The proposed approach integrates performance indicators, activity intensity measures, volatility metrics, and temporal lag features to capture both instantaneous and sequential learning behaviours. Rather than focusing solely on maximizing prediction accuracy, this study

emphasizes balanced classification behaviour and interpretability to support early performance tendency detection. The contributions of this study are threefold:

The development of a temporal, learning analytics–driven feature framework; a comprehensive comparison of individual and ensemble machine learning models; and the provision of interpretable insights into key behavioural factors influencing performance dynamics.

Literature Review

Learning Analytics in Digital Education

Learning analytics has emerged as a key research area in educational data mining, focusing on the systematic analysis of learner-generated data to understand and improve learning processes and outcomes [17]. Early studies in learning analytics emphasized descriptive and diagnostic approaches, such as monitoring learner engagement and summarizing academic performance [18]. With the widespread adoption of online learning platforms, learning analytics has progressively evolved toward predictive approaches that aim to anticipate learning outcomes and support timely interventions [19].

Several empirical studies demonstrate that learning analytics can effectively identify behavioural patterns related to student engagement, persistence, and academic success by leveraging interaction logs, assessment records, and activity traces from learning management systems [20], [21]. These findings underline the importance of behavioural and temporal features in modelling learning processes, particularly in environments characterized by continuous learner interaction.

Machine Learning for Student Performance Prediction

Machine learning techniques have been extensively applied to predict student performance and academic outcomes. Studies employing decision trees, neural networks, and other supervised learning models report promising predictive results when applied to structured educational datasets [22], [23]. However, performance is highly dependent on feature representation and data quality.

Recent research has shifted toward ensemble learning techniques, such as bagging, Random Forest, and Gradient Boosting, which aim to improve robustness and generalization capability [24], [25]. Empirical evidence suggests that ensemble-based models frequently outperform single classifiers in performance prediction tasks, particularly when dealing with heterogeneous and noisy educational data [26].

Temporal Modelling and Sequential Learning Behaviour

Learning is inherently a temporal process in which current performance is influenced by prior learning experiences and behavioural patterns. Consequently, temporal modelling has gained increasing attention in learning analytics research [26], [27]. Studies incorporating time-dependent features, such as lag variables and short-term behavioural changes, demonstrate improved predictive performance compared to static feature representations [28].

Sequential learning models, including recurrent neural networks and long short-term memory (LSTM) architectures, have been employed to capture complex

learning trajectories and engagement dynamics [23], [26]. Despite their strong representational power, such deep learning models often suffer from limited interpretability and high computational complexity, which restrict their practical adoption in educational settings [29].

Interpretability and Practical Adoption of AI in Learning

Interpretability is a critical requirement for AI-based systems in education, as predictive outcomes must be understandable and actionable for instructors and educational stakeholders [30]. Several studies emphasize that highly accurate but opaque models may hinder trust and limit real-world adoption.

To address this challenge, recent research has advocated for interpretable machine learning models and ensemble approaches combined with feature importance analysis [31]. Such methods provide a balance between predictive performance and transparency, enabling the identification of key behavioural factors influencing learning outcomes. Comprehensive surveys further highlight the need for interpretable, learning analytics–driven AI frameworks that support practical educational decision-making [32].

Research Methodology

This study adopts a quantitative research design using learning analytics–driven machine learning framework to predict short-term performance outcomes. As illustrated in [figure 1](#) (Research Workflow), the proposed methodology follows a structured pipeline that integrates data collection, preprocessing, feature engineering, model development, performance evaluation, and interpretability analysis. Historical learning-related data are first collected to capture performance indicators, activity intensity, and temporal variations. The preprocessing stage addresses data quality and consistency through missing value handling, categorical encoding, and normalization, ensuring that the dataset is suitable for supervised learning.

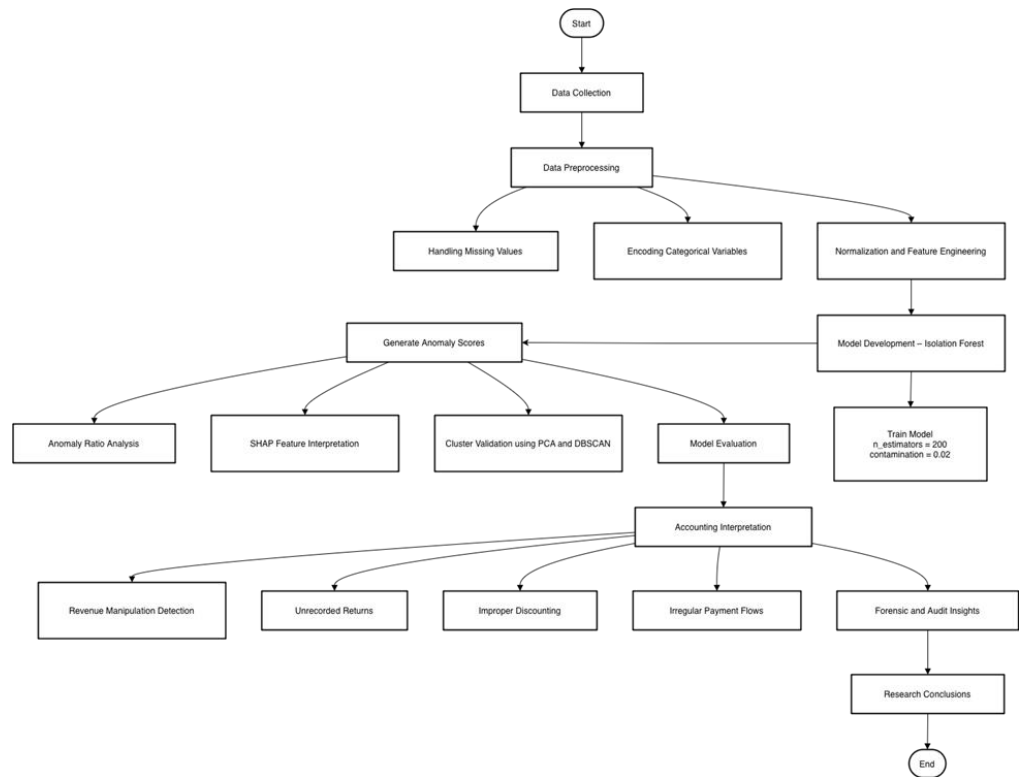


Figure 1 Research Workflow

Following preprocessing, informative learning analytics features are constructed to represent short-term performance changes, behavioural dynamics, and temporal dependencies. These features are used to train supervised classification models that distinguish between performance improvement and decline outcomes. Model performance is evaluated using multiple classification metrics to ensure balanced and reliable assessment. To enhance transparency and practical relevance, the workflow incorporates an interpretability stage that analyses feature importance and model behaviour, enabling the extraction of actionable insights. Overall, the research workflow presented in [figure 1](#) provides a systematic and reproducible framework that balances predictive capability and interpretability for AI-driven performance analysis in educational contexts.

Dataset and Data Preprocessing

The dataset used in this study consists of 1,334 temporal observations collected from a digital learning-related environment, where each observation represents a time-dependent performance snapshot. The dataset contains numerical indicators capturing multiple dimensions of learning behaviour, including performance states, activity intensity, and volatility-related characteristics. The temporal nature of the data enables the modelling of short-term performance tendencies by incorporating historical dependencies and sequential behavioural patterns.

Data preprocessing is conducted through a multi-stage procedure to ensure data quality and temporal integrity. First, the dataset is chronologically ordered based on its time index to preserve the sequential structure required for temporal analysis. Second, missing values arising from the construction of lag-

based features are removed to prevent information leakage and maintain data consistency across observations. Third, all variables are retained in their numerical form, as the selected machine learning models operate effectively on continuous-valued features and do not require categorical encoding in this context. Finally, the processed dataset is partitioned into training and testing subsets using a stratified splitting strategy, ensuring that the class distribution between performance improvement and decline outcomes remains balanced across both subsets. This approach supports robust model evaluation and reduces bias in performance assessment.

Feature Engineering

To capture both instantaneous and sequential learning behaviours, multiple categories of features are systematically constructed from the raw dataset. Performance indicators, including open, high, low, and close values, are used to represent different performance states within a given time interval, enabling the models to capture intra-period variation and performance range. Activity intensity is represented by Volume USDT, which serves as a proxy for the magnitude of learner engagement or interaction intensity within each time window.

Short-term performance dynamics are further modelled through a set of derived features designed to quantify relative change and volatility. These include return, which measures the proportional performance change between consecutive time steps; open_close_diff, which captures intra-interval performance movement; and high_low_spread, which reflects short-term volatility. To explicitly model temporal dependencies, lag-based features are incorporated, including close_lag1, close_lag2, and volume_lag1, allowing the models to learn from recent historical patterns and behavioural transitions. The target variable is defined as a binary performance outcome, indicating whether performance improves or declines in the subsequent time step. This formulation emphasizes early detection of short-term performance tendencies, rather than long-term outcome prediction, thereby aligning the modelling approach with real-time learning analytics and timely intervention objectives.

Machine Learning Models

Five machine learning models are employed in this study to evaluate predictive performance and robustness across different learning paradigms, while balancing interpretability and predictive capability.

The Decision Tree model is selected for its interpretability and ability to capture non-linear relationships between learning analytics features and performance outcomes. The model recursively partitions the feature space by selecting splits that maximize node purity. For classification tasks, impurity is measured using the Gini index, defined as:

$$Gini = 1 - \sum_{k=1}^K p_k^2 \quad (1)$$

p_k represents the proportion of samples belonging to class k in a given node.

Random Forest is an ensemble learning method based on bagging, where multiple decision trees are trained on bootstrapped subsets of the training data.

The final prediction is obtained through majority voting across all trees:

$$\hat{y} = \text{mode}\{h_1(x), h_2(x), \dots, h_T(x)\} \quad (2)$$

$h_t(x)$ denotes the prediction of the t -th decision tree and T is the total number of trees.

Gradient Boosting constructs models sequentially, where each new weak learner corrects the residual errors of the previous ensemble. At iteration m , the model update is expressed as:

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \quad (3)$$

$h_m(x)$ is the newly trained weak learner and γ_m denotes the learning rate.

XGBoost extends gradient boosting by incorporating regularization to control model complexity. The objective function optimized by XGBoost is given by:

$$\mathcal{L} = \sum_i l(y_i, \hat{y}_i) + \sum_k \Omega(f_k) \quad (4)$$

$l(y_i, \hat{y}_i)$ represents the loss function and $\Omega(f_k)$ is the regularization term.

Voting Ensemble

The Voting Ensemble combines predictions from multiple base classifiers using majority voting. For a binary classification task, the ensemble decision is defined as:

$$\hat{y} = \arg \max_{c \in \{0,1\}} \sum_{j=1}^N I(h_j(x) = c) \quad (5)$$

$I(\cdot)$ is the indicator function and N is the number of base classifiers.

The inclusion of both individual models and ensemble-based approaches enables a comprehensive comparison between interpretable classifiers and more complex ensemble techniques, allowing the study to assess trade-offs between predictive performance, robustness, and interpretability within learning analytics-driven framework.

Model Training and Evaluation

Model training is performed using the training subset of the dataset, while evaluation is conducted on a held-out testing subset to assess generalization performance. A stratified splitting strategy is applied to ensure that the class distribution of performance improvement and decline outcomes remains consistent across both subsets. This approach mitigates sampling bias and enables a fair comparison of predictive performance across models.

Model effectiveness is evaluated using four widely adopted classification metrics: Accuracy, Precision, Recall, and F1-score. Accuracy measures the overall proportion of correctly classified instances, while Precision and Recall provide class-sensitive evaluation by quantifying false positive and false negative behaviour, respectively. The F1-score, defined as the harmonic mean of Precision and Recall, is emphasized as the primary evaluation metric, as it

offers a balanced assessment of predictive performance, particularly in scenarios involving class imbalance or asymmetric misclassification costs. This metric is especially relevant for short-term performance prediction tasks, where both missed detections and false alarms may lead to suboptimal educational interventions.

To further examine model behaviour beyond aggregate metrics, confusion matrix analysis is employed, with particular focus on the Voting Ensemble model. The confusion matrix enables detailed inspection of true positive, true negative, false positive, and false negative predictions, thereby revealing class-wise prediction tendencies and potential bias toward specific outcome categories. This analysis provides deeper insight into how models respond to different performance states and supports a more nuanced interpretation of classification results.

Feature Importance and Interpretability Analysis

To enhance model transparency and support practical adoption, feature importance analysis is conducted using ensemble-based importance scores derived from tree-based models. This analysis quantifies the relative contribution of each learning analytics feature to the final prediction outcomes, enabling identification of the most influential indicators within the predictive framework.

Particular emphasis is placed on examining the roles of short-term performance change indicators, activity intensity measures, and temporal lag features, as these variables capture behavioural dynamics and recent learning history. By analysing feature importance rankings, the study provides empirical evidence on how recent performance fluctuations and engagement patterns influence subsequent performance tendencies.

Incorporating interpretability analysis ensures that the proposed framework aligns with the practical requirements of AI in learning environments, where predictive outputs must be explainable and actionable. The resulting insights enable educators and decision-makers to better understand underlying performance drivers, facilitating informed instructional adjustments and timely learning interventions.

Result

Dataset Characteristics and Feature Overview

This study utilizes a total of 1,334 observations derived from learning analytics-driven temporal dataset. [Table 1](#) presents the description of features employed in the predictive modelling process. The features include performance indicators, activity intensity measures, volatility metrics, and temporal lag variables, which collectively capture both instantaneous and sequential learning behaviours.

| Table 1 Dataset Description | | |
|-----------------------------|-------------------------------|---------|
| Feature | Description | Type |
| open | Opening performance indicator | Numeric |
| high | Maximum performance indicator | Numeric |

| | | |
|-----------------|---------------------------------|---------|
| low | Minimum performance indicator | Numeric |
| close | Closing performance indicator | Numeric |
| Volume USDT | Activity intensity indicator | Numeric |
| return | Performance change rate | Numeric |
| high_low_spread | Performance volatility | Numeric |
| open_close_diff | Intra-period performance change | Numeric |
| close_lag1 | Previous performance (t-1) | Numeric |
| close_lag2 | Previous performance (t-2) | Numeric |
| volume_lag1 | Lagged activity intensity | Numeric |

The distribution of performance outcome classes is shown in [Table 2](#). The dataset exhibits a nearly balanced class distribution, with 668 instances of performance improvement and 664 instances of performance decline. This balance reduces the risk of classification bias and supports reliable model evaluation.

| Table 2 Class Distribution of Performance Outcomes | | |
|--|---------|-------------------------|
| Class | Samples | Description |
| 1 | 668 | Performance Improvement |
| 0 | 664 | Performance Decline |

Temporal Performance Trend Analysis

The temporal evolution of the performance indicator is illustrated in [figure 2](#). The figure presents daily performance observations along with a 14-day moving average to smooth short-term fluctuations. Several periods of heightened volatility and regime shifts are observable throughout the timeline.

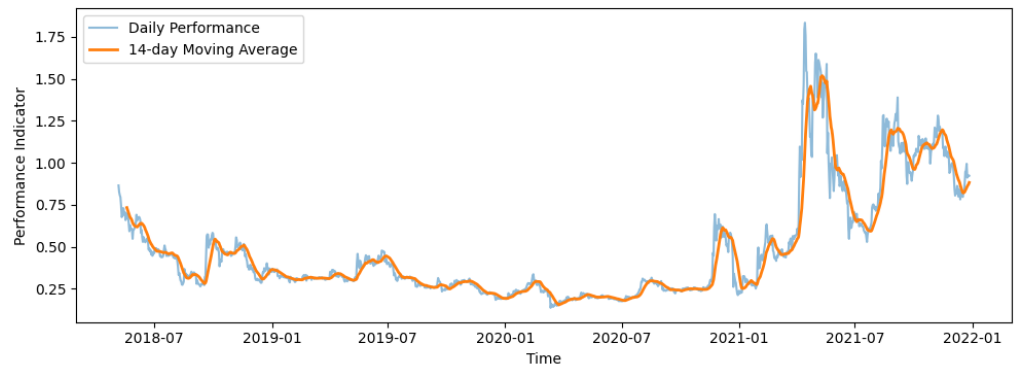


Figure 2 Performance trend over time with daily observations and a 14-day moving average smoothing

The pronounced variability observed in [figure 2](#) highlights the dynamic nature of performance trajectories and provides empirical justification for incorporating short-term change indicators and temporal lag features into the predictive framework.

Model Performance Comparison

The predictive performance of all machine learning models is summarized in [table 3](#) and visually compared in [figure 3](#). Evaluation metrics include Accuracy, Precision, Recall, and F1-score.

| Table 3 Model Performance Comparison | | | | |
|--------------------------------------|----------|-----------|--------|----------|
| Model | Accuracy | Precision | Recall | F1-score |
| Decision Tree | 0.539 | 0.540 | 0.552 | 0.546 |
| Random Forest | 0.513 | 0.515 | 0.515 | 0.515 |
| Gradient Boosting | 0.521 | 0.527 | 0.440 | 0.480 |
| XGBoost | 0.494 | 0.496 | 0.470 | 0.483 |
| Voting Ensemble | 0.510 | 0.505 | 0.505 | 0.502 |

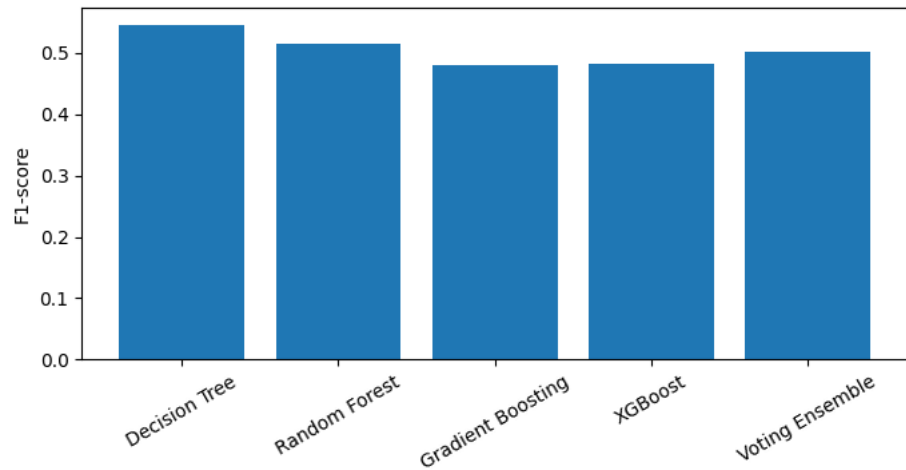


Figure 3 Model performance comparison using F1-score

Among individual classifiers, the Decision Tree model achieves the highest F1-score. The Voting Ensemble model provides more balanced performance across both classes, indicating improved stability in noisy temporal prediction tasks.

Confusion Matrix Analysis

To further examine classification behaviour, figure 4 presents the confusion matrix of the Voting Ensemble model.

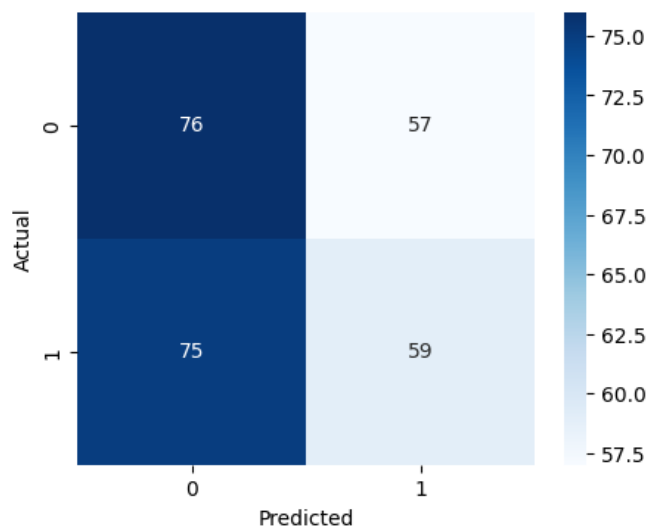


Figure 4 Confusion matrix of the Voting Ensemble model

The model correctly identifies 76 instances of performance decline and 59 instances of performance improvement. Although misclassifications are present, the distribution of errors remains relatively balanced, indicating no severe bias toward either outcome class.

Feature Importance Analysis

The contribution of individual features to model predictions is illustrated in [figure 5](#), which displays the top-ranked learning analytics indicators based on feature importance scores.

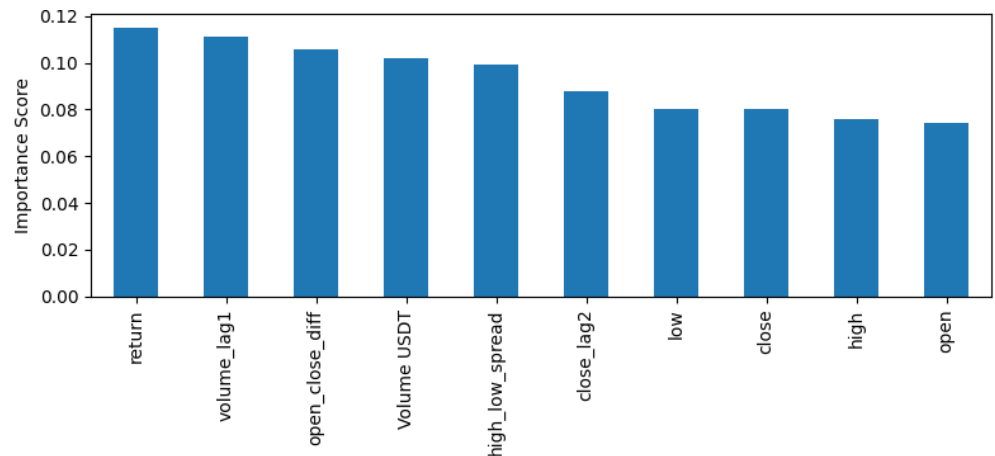


Figure 5 Top feature importance of learning analytics indicators

Short-term performance change measures, particularly return and open_close_diff, emerge as the most influential features. Lagged activity intensity and volatility-related features also play a significant role, emphasizing the importance of recent behavioural patterns and temporal dependencies in performance prediction.

Discussion

This study investigates the effectiveness of learning analytics-driven ensemble machine learning models for predicting short-term performance outcomes in a temporal learning context. The results demonstrate that, while overall predictive accuracy remains moderate, the proposed framework is capable of capturing meaningful behavioural patterns and performance tendencies from historical data.

The temporal analysis presented in [figure 2](#) highlights the highly dynamic nature of performance trajectories, characterized by frequent fluctuations and regime shifts. Such volatility inherently limits the predictability of next-step outcomes and explains the moderate performance observed across all models. This finding aligns with real-world learning environments, where performance is influenced not only by observable behavioural indicators but also by latent factors such as motivation, cognitive load, and external support that are not explicitly captured in system logs.

Model comparison results indicate that no single algorithm consistently dominates across all evaluation metrics. The Decision Tree model achieves the

highest F1-score among individual classifiers, suggesting that simpler models may effectively capture dominant patterns in short-term performance changes. In contrast, ensemble approaches such as Random Forest, Gradient Boosting, and XGBoost exhibit more conservative performance, reflecting their sensitivity to noise and rapid temporal transitions.

Although the Voting Ensemble model does not achieve the highest absolute F1-score, its balanced classification behaviour is particularly noteworthy. The confusion matrix analysis reveals comparable predictive capability for both performance improvement and decline classes, indicating reduced classification bias. In the context of learning analytics, such balanced detection is essential for early warning systems, where both declining and improving performance trends must be identified reliably to support timely interventions.

The feature importance analysis provides important interpretative insights into the underlying learning dynamics. The dominance of short-term performance change indicators, such as return and open_close_diff, suggests that recent outcome variations play a critical role in shaping subsequent performance. Additionally, the significance of lagged activity intensity and volatility-related features highlights the influence of sequential behavioural patterns and engagement intensity on learning outcomes. These findings reinforce the core premise of learning analytics, which emphasizes the value of temporal behavioural signals over static performance measures.

From an educational perspective, the proposed framework offers practical implications despite its moderate predictive accuracy. Rather than serving as a deterministic prediction tool, the model is better positioned as a performance tendency detection mechanism that can inform adaptive learning strategies and early intervention policies. By identifying periods of heightened risk or potential improvement, educational stakeholders can deploy targeted support measures before substantial performance changes occur.

Overall, the contribution of this study lies not in maximizing predictive accuracy, but in demonstrating how interpretable, learning analytics-driven AI models can uncover key behavioural factors influencing performance in dynamic learning environments. Future research may extend this work by incorporating richer contextual features, longer temporal dependencies, or deep learning architectures to further enhance predictive capability and robustness.

Conclusion

This study proposes learning analytics-driven framework for predicting short-term performance outcomes using ensemble machine learning models. By integrating performance indicators, activity intensity measures, volatility metrics, and temporal lag features, the proposed approach captures both instantaneous and sequential behavioural patterns in a temporal learning context.

Experimental results demonstrate that, although the predictive accuracy achieved by the models is moderate, meaningful performance tendencies can still be identified. The comparative analysis shows that individual models such as Decision Tree achieve competitive performance, while the Voting Ensemble model provides more balanced classification behaviour across performance improvement and decline classes. These findings highlight the importance of model stability and interpretability over solely maximizing accuracy in highly dynamic and noisy learning environments.

Feature importance analysis further reveals that short-term performance change indicators and recent activity intensity play a dominant role in shaping subsequent performance outcomes. This insight reinforces the central role of temporal learning analytics in understanding performance dynamics and supports the use of sequential behavioural features in predictive modelling.

From a practical perspective, the proposed framework is best positioned as a performance tendency detection tool rather than a deterministic prediction system. It can support early warning mechanisms and data-informed interventions by identifying potential performance risks and improvement opportunities at an early stage. Such capabilities are particularly valuable for adaptive learning systems and educational decision-making processes.

Future research may extend this work by incorporating richer contextual and behavioural features, exploring longer temporal dependencies, and applying advanced deep learning architectures to further enhance predictive robustness. Overall, this study contributes an interpretable and learning analytics-driven AI framework that advances the understanding of performance prediction in dynamic learning environments.

Declarations

Author Contributions

Conceptualization: E.D.L; Methodology: F.K.; Software: E.D.L; Validation: F.K.; Formal Analysis: E.D.L; Investigation: F.K.; Resources: F.K.; Data Curation: E.D.L; Writing Original Draft Preparation: F.K.; Writing Review and Editing: E.D.L; Visualization: F.K.; All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] G. Siemens and R. S. J. d. Baker, "Learning analytics and educational data mining: Towards communication and collaboration," in *Proceedings of the 2nd*

- International Conference on Learning Analytics and Knowledge (LAK)*, Vancouver, BC, Canada, vol. 2012, no. April, pp. 252–254, 2012, doi: 10.1145/2330601.2330661.
- [2] T. A. Dyckhoff, D. Zielke, M. Bültmann, M. Chatti, and U. Schroeder, “Design and implementation of a learning analytics toolkit for teachers,” *Educational Technology & Society*, vol. 15, no. 3, pp. 58–76, 2012, doi: 10.30191/ETS.201207_15(3).0006.
- [3] G. Siemens, “Learning analytics: Envisioning a research discipline and a domain of practice,” in *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK)*, vol. 2012, no. April, pp. 4–8, 2012, doi: 10.1145/2330601.2330605.
- [4] R. Ferguson, “Learning analytics: Drivers, developments and challenges,” *International Journal of Technology Enhanced Learning*, vol. 4, no. 5–6, pp. 304–317, 2012, doi: 10.1504/IJTEL.2012.051816.
- [5] S. D. Kotsiantis, C. J. Pierrakeas, and P. E. Pintelas, “Predicting students’ performance in distance learning using machine learning techniques,” *Applied Artificial Intelligence*, vol. 18, no. 5, pp. 411–426, 2004, doi: 10.1080/08839510490442058.
- [6] M. Cortez and A. Silva, “Using data mining to predict secondary school student performance,” *European Journal of Operational Research*, vol. 183, no. 1, pp. 49–61, 2008, doi: 10.1016/j.ejor.2006.10.038.
- [7] J. T. M. Tan, E. H. Lim, and C. F. Ng, “Predicting students’ academic performance using data mining techniques,” in *Proceedings of the IEEE International Conference on Advanced Learning Technologies (ICALT)*, pp. 102–104, 2011, doi: 10.1109/ICALT.2011.32.
- [8] R. S. Baker and K. Yacef, “The state of educational data mining in 2009: A review and future visions,” *Journal of Educational Data Mining*, vol. 1, no. 1, pp. 3–17, 2009, doi: 10.5281/zenodo.3554657.
- [9] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.
- [10] J. E. Beck and B. P. Woolf, “High-level student modelling with machine learning,” in *Proceedings of the 5th International Conference on Intelligent Tutoring Systems*, vol. 2000, no. January, pp. 584–593, 2000, doi: 10.1007/3-540-45108-0_62.
- [11] A. Hershkovitz and R. S. J. d. Baker, “Sequence mining in educational data,” *Journal of Educational Data Mining*, vol. 8, no. 1, pp. 1–20, 2016, doi: 10.5281/zenodo.3554650.
- [12] K. Piech et al., “Deep knowledge tracing,” *Advances in Neural Information Processing Systems*, vol. 28, pp. 505–513, 2015, doi: 10.5555/2969239.2969296.
- [13] C. Molnar, *Interpretable Machine Learning*, 2nd ed., Leanpub, 2022, doi: 10.5281/zenodo.4723246.
- [14] R. S. Baker, “Big data and education,” *Teachers College Record*, vol. 119, no. 9, pp. 1–26, 2017, doi: 10.1177/016146811711900901.
- [15] P. H. Winne and A. F. Hadwin, “Studying as self-regulated learning,” in *Metacognition in Educational Theory and Practice*, Routledge, pp. 277–304, 1998, doi: 10.4324/9780203876428.
- [16] Z. C. Lipton, “The mythos of model interpretability,” *Queue*, vol. 16, no. 3, pp. 30–

- 57, 2018, doi: 10.1145/3236386.3241340.
- [17] H. A. Mengash, "Using data mining techniques to predict student performance to support decision making in university admission systems," *IEEE Access*, vol. 8, no. March, pp. 55462–55470, 2020, doi: 10.1109/ACCESS.2020.2981905.
- [18] S. D. A. Bujang, A. Selamat, R. Ibrahim, O. Krejcar, H. Fujita, and E. Herrera-Viedma, "Multiclass prediction model for student grade prediction using machine learning," *IEEE Access*, vol. 9, no. June, pp. 95608–95621, 2021, doi: 10.1109/ACCESS.2021.3093563.
- [19] M. Yağcı, "Educational data mining: Prediction of students' academic performance using machine learning algorithms," *Smart Learning Environments*, vol. 9, no. 1, pp. 1–18, 2022, doi: 10.1186/s40561-022-00211-y.
- [20] A. Asif, A. Merceron, S. A. Ali, and N. G. Haider, "Analysing undergraduate students' performance using educational data mining," *Computers & Education*, vol. 113, no. October, pp. 177–194, 2017, doi: 10.1016/j.compedu.2017.05.007.
- [21] Z. Liu, G. Cen, Y. Wei, and Z. Li, "Multiple features fusion attention mechanism enhanced deep knowledge tracing for student performance prediction," *IEEE Access*, vol. 8, pp. 194894–194903, 2020, doi: 10.1109/ACCESS.2020.3033693.
- [22] Z. Chen, G. Cen, Y. Wei, and Z. Li, "Student performance prediction approach based on educational data mining," *IEEE Access*, vol. 11, no. November, pp. 131260–131272, 2023, doi: 10.1109/ACCESS.2023.3335985.
- [23] H. Waheed, S. U. Hassan, N. R. Aljohani, A. Alelyani, M. M. Almatrafi, and A. S. Alshehri, "Predicting the academic performance of students from VLE big data using deep learning models," *Computers in Human Behaviour*, vol. 104, no. March, p. 106189, 2020, doi: 10.1016/j.chb.2019.106189.
- [24] D. Injadat, A. Moubayed, A. B. Nassif, and A. Shami, "Multi-split optimized bagging ensemble model selection for multi-class educational data mining," *Applied Intelligence*, vol. 50, no. 12, pp. 4506–4528, 2020, doi: 10.1007/s10489-020-01806-4.
- [25] A. Asselman, M. Khaldi, and S. Aammou, "Enhancing the prediction of student performance based on the machine learning XGBoost algorithm," *Interactive Learning Environments*, vol. 31, no. 6, pp. 3360–3379, 2023, doi: 10.1080/10494820.2021.1928235.
- [26] Z. Xu, H. Yuan, and Q. Liu, "Student performance prediction based on blended learning," *IEEE Transactions on Education*, vol. 64, no. 1, pp. 66–73, 2021, doi: 10.1109/TE.2020.3048855.
- [27] R. Akour, M. Alsmadi, Y. Alazzam, and H. Aljamaan, "Predicting students' academic procrastination in blended learning courses using homework submission data," *IEEE Access*, vol. 7, pp. 102487–102498, 2019, doi: 10.1109/ACCESS.2019.2928569.
- [28] E. A. Amrieh, T. Hamtini, and I. Aljarah, "Mining educational data to predict student academic performance using ensemble methods," *International Journal of Database Theory and Application*, vol. 9, no. 8, pp. 119–136, 2016, doi: 10.14257/ijdta.2016.9.8.13.
- [29] A. Malik, E. M. Onyema, S. Dalal, U. K. Lilhore, and D. Anand, "Forecasting students' adaptability in online entrepreneurship education using modified ensemble machine learning model," *Array*, vol. 19, no. September, p. 100303, 2023, doi: 10.1016/j.array.2023.100303.

- [30] K. Latif, A. Malik, A. B. Nassif, and M. Alshazly, "Machine learning in higher education: Students' performance assessment considering online activity logs," *IEEE Access*, vol. 11, no. June, pp. 69586–69600, 2023, doi: 10.1109/ACCESS.2023.3287972.
- [31] Y. Lou, Y. Tang, J. Guo, and L. Zhang, "A systematic review on educational data mining methods used for predicting students' performance," *Engineering Reports*, vol. 4, no. 5, p. e12482, 2022, doi: 10.1002/eng2.12482.
- [32] C. Romero and S. Ventura, "Educational data mining and learning analytics: An updated survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 10, no. 3, p. e1355, 2020, doi: 10.1002/widm.1355.