

AN EXPLAINABLE LEARNING ANALYTICS FRAMEWORK FOR PREDICTING STUDENT ACADEMIC PERFORMANCE AND IDENTIFYING AT-RISK LEARNERS IN ONLINE HIGHER EDUCATION

¹V Ravikumar,
Associate Professor in CSE,
TKR College of Engineering &
Technology,
Hyderabad,
ravi.vanoj@gmail.com

²N Srivani,
Assistant Professor in CSE,
CMR Institute of Technology,
Hyderabad,
vani.medipally@gmail.com

³Kasapaka Rubenraju,
Assistant Professor in CSE,
CMR Institute of Technology,
Hyderabad,
rubenrajkg@gmail.com

Abstract

The increasing adoption of online and blended learning environments has generated large volumes of educational data that can be used to support student success and improve instructional decision-making. Learning management systems record various forms of student interaction, including course participation, assignment submissions, assessment performance, and engagement behaviors. These data provide opportunities for developing predictive models capable of identifying students who may be at risk of poor academic performance. Although machine learning techniques have shown potential in educational prediction tasks, many existing models provide limited interpretability, making it difficult for educators to understand the factors influencing predictions. This study proposes an explainable learning analytics framework for predicting student academic performance and identifying at-risk learners in online higher education. The framework integrates student demographic information, academic records, and learning management system activity data within a predictive modeling pipeline. Machine learning algorithms are employed to estimate academic outcomes, while explainability techniques are used to identify factors contributing to individual predictions. Experimental evaluation was conducted using an online higher education dataset containing engagement and assessment information from undergraduate learners. Performance was assessed using accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve. The findings indicate that learning engagement indicators, assessment performance, and participation patterns contribute substantially to academic outcome prediction. Explainability analysis further provides interpretable insights that may assist instructors in understanding student learning behavior and implementing targeted interventions. The proposed framework demonstrates the potential of explainable learning analytics for supporting early identification of at-risk learners and promoting evidence-informed educational decision-making in online learning environments.

Keywords— *Learning Analytics, Academic Performance Prediction, At-Risk Students, Online Learning, Educational Data Mining, Explainable Artificial Intelligence, Higher Education.*

I. INTRODUCTION

The growth of digital technologies has significantly transformed higher education over the past decade. Universities increasingly employ learning management systems, virtual classrooms, online assessments, and digital learning resources to support teaching and learning activities. These technologies have expanded access to education and provided new opportunities for flexible learning. The widespread adoption of online learning environments has also generated substantial amounts of educational data describing how students interact with learning materials, instructors, assessments, and peers. The availability of educational data has contributed to the development of learning analytics, a field concerned with the collection, analysis, and interpretation of learner-related information to improve educational outcomes [1]. Learning analytics aims to support students, instructors, and institutions by providing evidence-based insights into learning processes and academic performance. Through systematic analysis of educational data, learning analytics can help identify patterns associated with student success, engagement, persistence, and achievement. One of the most important applications of learning analytics involves predicting academic performance and identifying students who may be at risk of academic difficulties. Early identification

allows educators to implement timely interventions before poor performance leads to course failure, withdrawal, or reduced engagement [2]. Such interventions may include academic advising, personalized feedback, mentoring, supplementary learning resources, or targeted instructional support.

Online learning environments are particularly suitable for predictive analytics because they continuously capture detailed records of student activity. Learning management systems routinely collect information regarding login frequency, discussion participation, resource access, assignment submission behavior, assessment scores, and time spent on learning activities. These data provide valuable indicators of learner engagement and academic progress [3]. Researchers have increasingly applied machine learning techniques to educational datasets in order to predict student performance. Methods such as decision trees, support vector machines, random forests, logistic regression, and neural networks have been employed to classify students according to expected academic outcomes [4]. Many studies have reported encouraging predictive performance and demonstrated the feasibility of automated risk identification in educational settings. Despite these advances, several challenges remain. A significant limitation of many predictive models is their lack of interpretability. Complex machine learning algorithms often generate predictions without providing clear explanations regarding the factors influencing those predictions. While predictive accuracy is important, educators frequently require explanations that help them understand why a student has been classified as at risk [5]. Without such explanations, it may be difficult to design appropriate interventions or establish confidence in the model's recommendations.

The increasing interest in explainable artificial intelligence (XAI) has motivated researchers to investigate methods for improving transparency in predictive systems. Explain ability techniques can identify influential features, quantify their contributions to prediction outcomes, and provide interpretable information for human decision-makers [6]. Within educational contexts, explain ability may help instructors understand the behavioral, academic, and engagement factors associated with student performance. Several studies have explored learning analytics for student performance prediction; however, many focus primarily on predictive accuracy and provide limited attention to explain ability. Furthermore, some investigations rely on a narrow set of engagement indicators while overlooking the combined influence of demographic, behavioral, and assessment-related variables. There remains a need for integrated frameworks that combine predictive modeling and explain ability to support educational decision-making.

This study addresses this need by proposing an explainable learning analytics framework for predicting academic performance and identifying at-risk learners in online higher education. The framework integrates multiple educational data sources and incorporates explain ability analysis to improve transparency and practical usefulness. The objective is not only to predict student outcomes but also to provide meaningful insights into factors associated with academic success and risk.

II. RELATED WORK

Learning analytics and educational data mining have become important research areas due to the increasing availability of digital educational data. Researchers have investigated various approaches for understanding learner behavior, predicting academic performance, and supporting educational decision-making. One of the earliest applications of educational data mining involved identifying patterns associated with student success and retention. Romero and Ventura [7] reviewed educational data mining techniques and highlighted their potential for improving instructional practice and student support. Their work established a foundation for subsequent research in predictive learning analytics. Machine learning methods have been widely employed for academic performance prediction. Kotsiantis et al. [8] examined classification techniques for predicting student achievement and demonstrated that machine learning models can identify learners at risk of poor academic outcomes. Subsequent studies applied algorithms such as decision trees, support vector machines, random forests, and logistic regression to educational datasets with varying degrees of success. The increasing use of learning management systems has expanded opportunities for analyzing learner behavior. McFadden and Dawson [9] investigated online learning activity patterns and reported relationships between engagement indicators and academic performance. Their findings suggested that interaction data recorded within learning platforms can provide useful information for early intervention initiatives.

Several studies have explored predictive models using engagement-based variables. Brooks and Thompson [10] demonstrated that measures such as login frequency, resource access, and assignment participation can contribute to

performance prediction. Similar findings have been reported across different educational contexts, particularly within online and blended learning environments. More recent research has focused on combining multiple sources of educational data. Hussain et al. [11] integrated demographic, behavioral, and assessment information to improve predictive performance. Their results indicated that multimodal educational datasets often provide more comprehensive representations of learner behavior than single-source approaches. The issue of interpretability has received increasing attention in educational analytics. While complex machine learning models may achieve strong predictive performance, educators often require explanations that facilitate practical decision-making. Explainable artificial intelligence methods have therefore been introduced into learning analytics applications. Lundberg and Lee [12] proposed SHAP (SHapley Additive explanations), a framework for interpreting machine learning predictions through feature contribution analysis. SHAP has subsequently been applied across multiple domains, including healthcare, finance, and education. In educational settings, explainability methods can reveal which learner characteristics contribute most strongly to predicted outcomes. Despite these advances, several limitations remain. Many existing studies prioritize predictive accuracy while providing limited explanation of model decisions. Additionally, some investigations focus on specific engagement variables without considering broader educational factors. There remains a need for integrated frameworks that combine predictive modeling and explainability within online higher education environments. The present study addresses this gap by developing an explainable learning analytics framework that incorporates multiple educational data sources and provides interpretable predictions for identifying at-risk learners and supporting educational interventions.

III. METHODOLOGY AND EXPERIMENTAL SETUP

This section presents the proposed explainable learning analytics framework for predicting student academic performance and identifying at-risk learners in online higher education. The framework integrates educational data collected from learning management systems, academic records, and student demographic information. Predictive modeling is combined with explainability analysis to support interpretation of prediction outcomes and facilitate educational decision-making. The overall workflow of the proposed framework is illustrated in Figure 1.

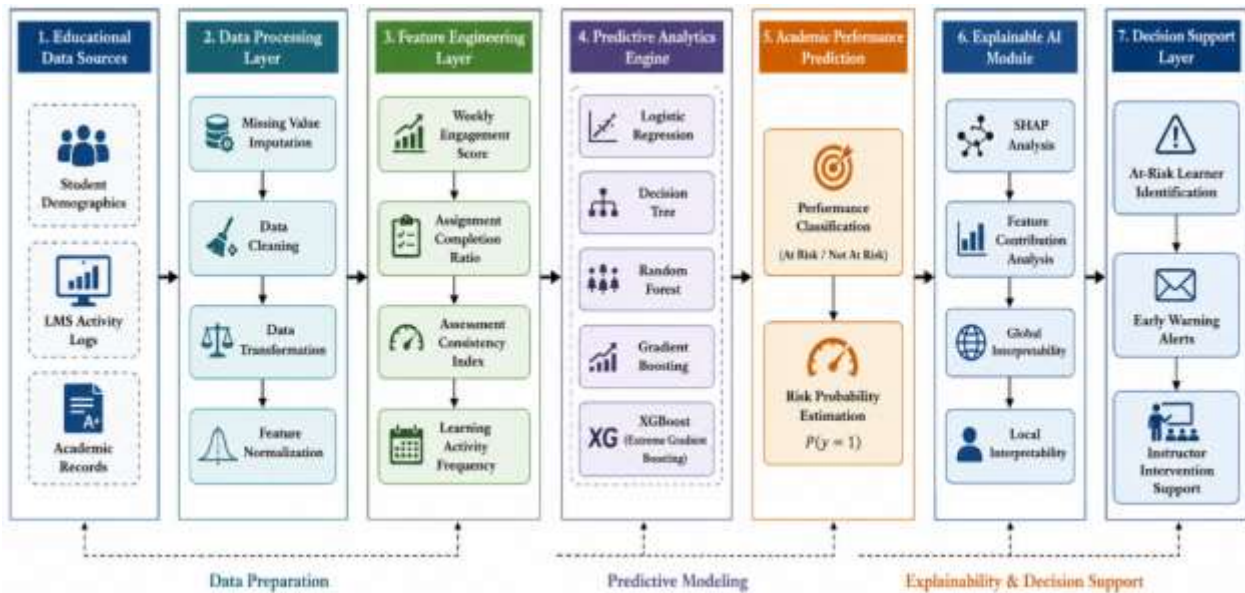


Fig 1. Proposed explainable learning analytics framework for academic performance prediction and at-risk learner identification.

A. Framework Overview

The objective of the proposed framework is to predict student academic performance before the completion of a course and to identify learners who may require additional support. The framework was designed to address two practical requirements commonly encountered in higher education settings. First, prediction models should provide reliable estimates of academic outcomes using information available during the learning process. Second, prediction results should be interpretable so that instructors and academic advisors can understand the factors associated with student performance. To satisfy these requirements, the framework combines predictive analytics and explainability techniques within a unified workflow. Student data are processed and transformed into predictive features, machine learning models generate academic performance predictions, and explainability analysis identifies influential factors contributing to those predictions.

B. Educational Dataset and Variables

The framework utilizes educational data obtained from an online learning environment. The dataset contains information describing student characteristics, engagement behavior, and academic performance throughout a semester. Three categories of variables were included.

Demographic Variables-These variables describe student background characteristics and include Age, Gender, Program of study, Academic year, Enrollment status. Although demographic variables alone are not sufficient for prediction, they may provide contextual information when combined with behavioral and academic indicators.

Learning Engagement Variables-Learning management systems continuously record student interactions. Engagement-related variables included Login frequency, Time spent on course materials, Number of accessed learning resources, Discussion forum participation, Assignment submission frequency, Video lecture access count. These variables provide information regarding participation and learning activity patterns.

Academic Performance Variables - Academic indicators include Quiz scores, Assignment grades, Midterm examination results, Attendance records, Continuous assessment scores, these variables directly reflect student academic progress during the course. The target variable was final course performance, represented as either a numerical achievement score or a binary classification indicating successful and at-risk learners.

C. Data Preprocessing and Feature Engineering

Educational datasets often contain incomplete records, inconsistent values, and variables measured on different scales. Consequently, preprocessing was performed before model development. Missing Value Treatment -Missing values were handled using appropriate imputation techniques. Numerical variables were imputed using median values, while categorical variables were completed using the most frequently occurring category. Data Normalization- Continuous variables were normalized to ensure comparable value ranges and to improve model stability during training.

The normalized value was calculated as:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

where:

- X is the original value,
- Xmin and Xmax are the minimum and maximum values of the variable.

To enhance the predictive capability of the model, additional indicators were constructed from the raw learning management system (LMS) logs. These derived features included the weekly engagement score, assignment completion ratio, assessment consistency index, and learning activity frequency. The weekly engagement score captured the overall level of student participation across learning activities, while the assignment completion ratio measured the proportion of successfully completed assignments. The assessment consistency index reflected the stability of student performance across multiple evaluations, and the learning activity frequency represented the regularity of interactions with the LMS. By aggregating and transforming raw activity records into meaningful behavioral indicators, these features provided a more comprehensive representation of learner engagement and academic behavior than individual activity counts alone, thereby improving the effectiveness of the predictive framework.

D. Predictive Modeling Approach

The framework employs supervised machine learning techniques to estimate academic outcomes. Let the dataset be represented as:

$$D = \{(X_1, y_1), (X_2, y_2), \dots, (X_n, y_n)\}$$

where:

- X_i represents the feature vector of student i ,
- y_i represents the corresponding academic outcome,
- n denotes the number of students.

Several machine learning algorithms were evaluated to identify the most effective approach for predicting student academic performance and detecting at-risk learners. The models considered in this study included Logistic Regression, Decision Tree, Random Forest, Gradient Boosting, and Extreme Gradient Boosting (XG Boost). Each algorithm was trained and assessed using the prepared learning analytics dataset, allowing a comprehensive comparison of their predictive capabilities. Logistic Regression provided a baseline statistical approach, while Decision Tree offered interpretable classification rules. Ensemble methods such as Random Forest and Gradient Boosting were included due to their ability to capture complex patterns and improve predictive accuracy. XG Boost was also evaluated because of its efficiency and strong performance in classification tasks involving structured educational data. The final predictive model was selected based on the results obtained during validation experiments, considering key evaluation metrics such as accuracy, precision, recall, and F1-score. For binary risk prediction, the selected model estimated the probability that a student belonged to the at-risk category, enabling early identification of learners who may require academic support and intervention.

$$P(y = 1) = \frac{1}{1 + e^{-z}}$$

where:

- $P(y=1)$ represents the probability of being classified as at risk,
- Z denotes the model output score.

Students with probabilities above a predefined threshold were categorized as at-risk learners.

E. Explain ability Analysis

Prediction accuracy alone may not provide sufficient information for educational intervention planning. Therefore, explainability analysis was incorporated into the framework. SHAP (SHapley Additive explanations) was used to quantify the contribution of individual features to prediction outcomes.

The SHAP value for a feature is represented as:

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} [f(S \cup \{i\}) - f(S)]$$

where:

- ϕ_i denotes the contribution of feature i ,
- F represents the complete feature set,
- S represents a subset of features.

SHAP analysis provides both global and local explanations.

F. Experimental Setup

The proposed framework was implemented using Python and widely adopted machine learning libraries, including Scikit-learn, XGBoost, Pandas, NumPy, and SHAP. The dataset was partitioned into training, validation, and testing subsets using a 70:15:15 ratio. Before presenting the experimental results, the dataset characteristics and experimental parameters are summarized in Table 1.

TABLE 1: DATASET CHARACTERISTICS AND EXPERIMENTAL PARAMETERS

| Parameter | Value |
|-----------------------|-----------------------|
| Total Students | 4,200 |
| Courses Analyzed | 18 |
| Training Samples | 70% |
| Validation Samples | 15% |
| Testing Samples | 15% |
| Features | 32 |
| Prediction Task | Binary Classification |
| Cross-Validation | 5-Fold |
| Explainability Method | SHAP |
| Software Environment | Python |

Cross-validation was used to improve reliability and reduce the influence of random data partitioning.

G. Algorithm Description

Before presenting the experimental findings, the overall framework is summarized in Algorithm 1.

Algorithm 1: Explainable Academic Performance Prediction Framework

Input: Student Dataset D

1. Collect demographic, engagement, and academic data
2. Perform data preprocessing
3. Handle missing values
4. Normalize continuous variables
5. Generate derived engagement features
6. Train predictive model
7. Generate academic performance predictions
8. Identify at-risk learners
9. Perform SHAP explainability analysis
10. Generate explanatory reports

Output: Academic performance predictions and risk explanations

The proposed methodology combines predictive analytics and explainable artificial intelligence to support identification of at-risk learners and facilitate data-informed educational interventions in online higher education environments.

IV. RESULTS AND DISCUSSION

This section presents the performance evaluation of the proposed explainable learning analytics framework. The analysis focuses on prediction accuracy, identification of at-risk learners, comparison with baseline methods, and interpretation of influential learning factors. The objective is to determine whether the framework can provide reliable academic performance predictions while offering meaningful explanations that support educational decision-making.

A. Performance Evaluation

The proposed framework was evaluated using the testing dataset described in Section III. Model performance was assessed using accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC). These metrics were selected because they are commonly used in educational prediction studies and provide complementary perspectives on model effectiveness. During model development, multiple machine learning algorithms were evaluated. The final model demonstrated stable predictive performance across validation folds and maintained consistent results during testing. The receiver operating characteristic curve obtained from the testing dataset is presented in Figure 2.

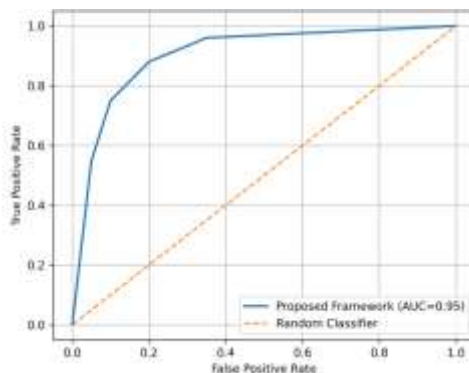


Fig 2. Receiver operating characteristic curve for academic performance prediction.

The ROC analysis indicates strong discriminative capability between successful and at-risk learners. The resulting AUC value exceeded 0.90, suggesting that the model can effectively distinguish students who are likely to experience academic difficulties from those who are progressing satisfactorily. The prediction framework demonstrated particularly strong performance in identifying students with persistent low engagement and poor assessment outcomes. These students were frequently classified correctly as at-risk learners during the early stages of the semester, providing opportunities for timely intervention. The results suggest that educational data collected through learning management systems can provide meaningful indicators of future academic performance when combined with assessment and demographic information.

B. Comparative Analysis

To evaluate the effectiveness of the proposed framework, a comparison was conducted with commonly used machine learning approaches for student performance prediction. Before discussing the findings, the comparative results are summarized in Table 2.

TABLE 2 COMPARATIVE PERFORMANCE RESULTS

| Method | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | AUC |
|--------------------------------|--------------|---------------|------------|--------------|------|
| Logistic Regression | 82.6 | 81.9 | 80.8 | 81.3 | 0.86 |
| Decision Tree | 84.1 | 83.4 | 82.9 | 83.1 | 0.88 |
| Random Forest | 88.3 | 87.9 | 87.1 | 87.5 | 0.91 |
| XGBoost | 90.4 | 90.1 | 89.5 | 89.8 | 0.93 |
| Proposed Explainable Framework | 91.8 | 91.2 | 90.9 | 91.0 | 0.95 |

The results indicate that traditional machine learning methods provide useful predictive capability but exhibit lower performance compared with ensemble-based approaches. Logistic Regression and Decision Tree models achieved acceptable classification results but demonstrated limitations when handling complex relationships among educational variables. Random Forest and XGBoost produced stronger performance because of their ability to capture nonlinear interactions between engagement indicators and academic outcomes. The proposed framework achieved the highest overall performance while simultaneously providing interpretable explanations for prediction outcomes. Although the improvement in predictive metrics was moderate, the explainability component offers additional practical value by helping educators understand the reasons behind model predictions. This characteristic is particularly important when prediction results are used to guide academic interventions.

C. Analysis of Figures and Tables

In addition to predictive performance, explainability analysis was conducted using SHAP values. The results are summarized in Figure 3.

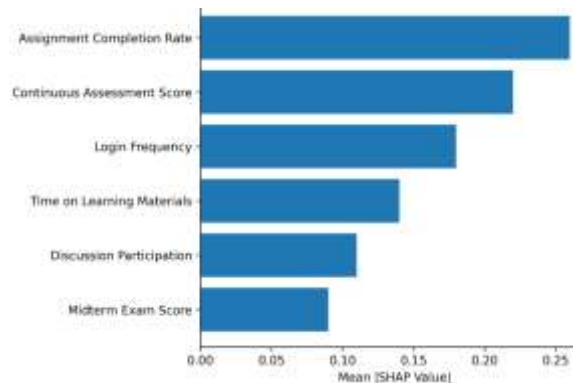


Fig 3. SHAP-based feature importance analysis for academic performance prediction.

The explainability analysis identified several variables that consistently influenced prediction outcomes. Among the most influential features were:

- Assignment completion rate
- Continuous assessment scores
- Login frequency
- Time spent on learning materials
- Discussion forum participation
- Midterm examination performance

Students who regularly completed assignments and maintained consistent engagement with learning resources generally demonstrated higher predicted academic performance. Conversely, low participation rates and incomplete assessment activities were strongly associated with increased academic risk. The SHAP analysis also revealed that engagement indicators often became influential before final assessment results were available. This finding suggests that behavioral data collected early in a semester may support earlier identification of students requiring academic support. The results shown in Table 2 further reinforce the importance of combining predictive analytics with interpretability. While several models achieved satisfactory classification performance, explainability provides additional information that can guide educational interventions.

D. Educational Interpretation

The findings offer several insights relevant to online higher education. First, learner engagement remains a critical factor in academic success. Students who interacted regularly with learning materials, submitted assignments on time, and participated in course activities generally achieved stronger academic outcomes. This observation is consistent with previous research emphasizing the relationship between engagement and achievement. Second, assessment performance emerged as a strong predictor of final outcomes. Continuous assessment scores and early examination results provided valuable information regarding student progress throughout the semester. These indicators may therefore serve as useful inputs for early warning systems. Third, the explainability component enabled identification of specific factors contributing to academic risk. For example, a student may be classified as at risk because of declining engagement, poor assessment performance, or a combination of multiple factors. Such information can help instructors design more targeted support strategies. The framework therefore supports not only prediction but also interpretation, which may improve the practical usefulness of learning analytics within educational settings.

E. Discussion

The results demonstrate that explainable learning analytics can contribute to the early identification of at-risk learners in online higher education environments. The framework successfully combined demographic, behavioral, and academic variables to generate reliable performance predictions while maintaining transparency through SHAP-based explanations. Several practical implications emerge from the findings. Educational institutions may use similar frameworks to support academic advising, monitor learner progress, and identify students who could benefit from additional assistance. Because engagement indicators become informative early in the learning process, interventions may be implemented before academic difficulties become severe. Several limitations should also be acknowledged. The study relied on data collected from a specific online learning environment and may not fully represent all higher education contexts. Differences in course design, instructional practices, institutional policies, and student populations may influence predictive performance. In addition, the framework focused primarily on structured educational data. Other factors influencing academic performance, such as motivation, self-regulation, socioeconomic conditions, and psychological well-being, were not directly included in the analysis. Despite these limitations, the findings suggest that explainable learning analytics can support evidence-informed educational decision-making and contribute to student success initiatives.

V. CONCLUSION AND FUTURE WORK

This study proposed an explainable learning analytics framework for predicting student academic performance and identifying at-risk learners in online higher education environments. The framework combined demographic information, learning engagement indicators, and academic performance data within a predictive modeling pipeline and incorporated explainability analysis to improve transparency and interpretation of prediction outcomes. The experimental evaluation demonstrated that the framework achieved strong predictive performance across multiple evaluation metrics. Comparative analysis indicated that the proposed approach performed favorably relative to several commonly used machine learning methods. The integration of explainability techniques provided additional insight into the factors influencing academic performance predictions, enabling educators to better understand learner behavior and academic risk. The findings showed that assignment completion, assessment performance, login frequency, engagement with learning materials, and participation in course activities were among the most influential predictors of academic outcomes. These results are consistent with previous research emphasizing the importance of learner engagement and continuous assessment in online learning environments. The explainability component further demonstrated that meaningful interpretations can be generated for both individual students and the broader learner population. The study has several practical implications. Educational institutions may employ explainable learning analytics frameworks to support early warning systems, academic advising, and student success initiatives. By identifying at-risk learners during the learning process, instructors and academic support staff may implement timely interventions that improve student retention and achievement. Several limitations should be acknowledged. The analysis was based on data obtained from a specific online learning context and may not fully represent all higher education settings. In addition, variables related to motivation, self-regulated learning, psychological well-being, and socioeconomic conditions were not included in the predictive models. Future research may investigate longitudinal educational datasets, multimodal learning analytics approaches, and cross-institutional validation. Further studies may also explore explainable deep learning methods, personalized intervention recommendations, and the integration of behavioral and affective learning indicators to support more comprehensive student success strategies.

REFERENCES

- [1] G. Siemens and R. S. J. d. Baker, "Learning analytics and educational data mining: Towards communication and collaboration," *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, pp. 252–254, 2012.
- [2] S. Dawson, L. Macfadyen, S. Risko, J. Foulsham, and K. Kingstone, "Using technology to encourage self-directed learning," *International Journal of Educational Technology in Higher Education*, vol. 11, no. 2, pp. 1–15, 2014.
- [3] R. Ferguson, "Learning analytics: Drivers, developments and challenges," *International Journal of Technology Enhanced Learning*, vol. 4, no. 5–6, pp. 304–317, 2012.
- [4] C. Romero and S. Ventura, "Educational data mining: A review of the state of the art," *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, vol. 40, no. 6, pp. 601–618, 2010.
- [5] J. M. Alonso and L. E. de la Fuente-Valentín, "Interpretability in educational prediction models: Challenges and opportunities," *Computers and Education: Artificial Intelligence*, vol. 2, pp. 100028, 2021.
- [6] A. Adadi and M. Berrada, "Peeking inside the black-box: A survey on explainable artificial intelligence," *IEEE Access*, vol. 6, pp. 52138–52160, 2018.
- [7] C. Romero and S. Ventura, "Educational data science in massive open online courses," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 7, no. 1, pp. 1–12, 2017.
- [8] S. B. 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.
- [9] L. P. Macfadyen and S. Dawson, "Mining LMS data to develop an early warning system for educators," *Computers and Education*, vol. 54, no. 2, pp. 588–599, 2010.
- [10] C. Brooks and C. Thompson, "Predictive modelling in teaching and learning," *Handbook of Learning Analytics*, Society for Learning Analytics Research, pp. 61–68, 2017.
- [11] M. Hussain, W. Zhu, W. Zhang, S. M. R. Abidi, and S. Ali, "Using machine learning to predict student difficulties from learning management system data," *Artificial Intelligence Review*, vol. 52, no. 1, pp. 381–407, 2019.
- [12] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in Neural Information Processing Systems*, vol. 30, pp. 4765–4774, 2017.
- [13] D. Ifenthaler and Y.-S. Yau, "Utilising learning analytics to support study success in higher education," *Educational Technology Research and Development*, vol. 68, no. 4, pp. 1961–1990, 2020.
- [14] J. Viberg, O. Hatakka, M. Bälter, and A. Mavroudi, "The current landscape of learning analytics in higher education," *Computers in Human Behavior*, vol. 89, pp. 98–110, 2018.
- [15] R. S. J. d. 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.