e-ISSN: 3108-9828 | p-ISSN: 3109-3981
Volume 3 No 1 May 2026 | Page: 10–19
DOI: https://doi.org/10.64803/jocsaic.v3i1.174

Journal of Computer Science Artificial Intelligence and Communications

Journal Homepage: https://journals.raskhamedia.or.id/index.php/jocsaic

Explainable Artificial Intelligence for Student Academic Performance Prediction Using Random Forest and SHAP

Wahyu Nugraha[1*], Rabiatus Sa'adah[2], Muhammad Hasanuddin[3]

[1,2]Fakultas Teknik dan Informatika, Program Studi Informatika, Universitas Bina Sarana Informatika, Pontianak, Indonesia
[3]Magister Teknologi Informasi, Universitas Pembangunan Panca Budi, Medan, Indonesia
Author(s) Email: [1*]wahyu.whn@bsi.ac.id, [2]rabiatus.rbh@bsi.ac.id, [3]muhammadhasan20feb@gmail.com

ARTICLE INFO

Article history:

Received May 17, 2026
Revised May 25, 2026
Accepted May 29, 2026
Publish May 31, 2026

ABSTRACT

Predicting student academic performance has become an important application of educational data mining because it enables educational institutions to identify students who require academic support at an early stage. Although machine learning algorithms have demonstrated high predictive capability, many predictive models operate as black-box systems, making it difficult for educators to understand the factors influencing prediction outcomes. This study proposes an Explainable Artificial Intelligence (XAI) framework for student academic performance prediction by integrating the Random Forest algorithm with SHapley Additive exPlanations (SHAP). The proposed methodology consists of data collection, data preprocessing, feature selection, model development, performance evaluation, and model interpretation. Random Forest was employed as the primary classification algorithm due to its robustness and high predictive performance, while SHAP was utilized to provide transparent explanations of both global and local prediction results. The experimental evaluation demonstrated that the proposed model achieved high classification performance, obtaining an accuracy of 91.67%, precision of 90.32%, recall of 93.33%, and an F1-score of 91.80%. Furthermore, SHAP analysis identified Previous GPA, Final Examination Score, Attendance, Assignment Score, and Study Hours as the most influential factors affecting student academic performance. The integration of Random Forest and SHAP not only improves prediction reliability but also enhances model transparency by explaining the contribution of each feature to prediction outcomes. Consequently, the proposed framework supports evidence-based academic decision-making, facilitates early identification of at-risk students, and provides educators with interpretable insights for designing effective academic intervention strategies. These findings demonstrate that Explainable Artificial Intelligence can significantly improve the practical applicability and trustworthiness of machine learning models in educational environments.

Keywords:

Explainable Artificial Intelligence, Random Forest, SHAP, Student Academic Performance, Educational Data Mining.

Corresponding Author:
Wahyu Nugraha,
Fakultas Teknik dan Informatika, Program Studi Informatika, Universitas Bina Sarana Informatika, Pontianak, Indonesia
Email: wahyu.whn@bsi.ac.id

1. INTRODUCTION

The rapid advancement of digital technologies has significantly transformed the education sector by enabling institutions to collect and analyze large volumes of academic data[1],[2]. Educational institutions increasingly rely on data-driven decision-making to improve learning quality, monitor student progress, and provide timely academic support[3]. One of the most important applications of educational data mining is predicting student academic performance, as accurate predictions enable educators to identify students who may experience academic difficulties before their performance declines[4]. Early prediction allows institutions to implement appropriate interventions, improve student retention, and optimize academic achievement.

Machine learning has become one of the most widely adopted approaches for predicting student academic performance because of its ability to discover complex relationships among academic, behavioral, and demographic variables[5]. Various algorithms, including Decision Tree, Support Vector Machine, Artificial Neural Network, Naïve Bayes, Gradient Boosting, and Random Forest, have been successfully applied to educational datasets[6]. Among these algorithms, Random Forest has consistently demonstrated strong predictive capability due to its ensemble learning mechanism, robustness against overfitting, and ability to process datasets containing numerous attributes[7]. Furthermore, Random Forest requires relatively limited parameter tuning while maintaining stable performance across various educational datasets, making it a practical choice for academic performance prediction[8],[9].

Despite the promising predictive accuracy achieved by many machine learning algorithms, most predictive models function as black-box systems[10]. Although these models generate reliable predictions, they often fail to provide understandable explanations regarding how individual variables influence prediction outcomes[11]. In educational settings, this lack of interpretability presents a significant limitation because educators, academic advisors, and institutional decision-makers require transparent information before implementing intervention strategies. Understanding why a student is predicted to achieve a certain academic outcome is equally important as obtaining the prediction itself. Consequently, interpretability has become a critical requirement in educational artificial intelligence applications.

Explainable Artificial Intelligence (XAI) has recently emerged as an important research area that seeks to improve the transparency and interpretability of machine learning models without sacrificing predictive performance[12]. XAI enables users to understand the reasoning behind model decisions by revealing how input features contribute to prediction results. Rather than treating machine learning models as inaccessible black boxes, XAI transforms them into interpretable systems capable of supporting trustworthy and accountable decision-making. This capability is particularly valuable in education, where prediction outcomes may directly influence academic policies and student support programs.

Among various XAI techniques, SHapley Additive exPlanations (SHAP) has gained considerable attention because it provides theoretically sound and consistent explanations based on cooperative game theory. SHAP quantifies the contribution of each feature to individual predictions while simultaneously providing global interpretations of model behavior. Through visualization techniques such as summary plots, dependence plots, force plots, and feature importance analysis, SHAP enables researchers and educators to identify which variables most strongly influence academic performance[13],. These explanations improve confidence in predictive models and facilitate evidence-based academic decision-making.

Recent studies have demonstrated the effectiveness of machine learning in predicting student academic performance. However, many previous studies primarily focused on maximizing prediction accuracy while providing limited discussion regarding model interpretability. Although several studies incorporated feature importance analysis, traditional importance measures often fail to explain the contribution of individual features for specific prediction instances. Consequently, educators may still encounter difficulties when interpreting model decisions and translating prediction results into practical intervention strategies. This gap indicates the necessity of integrating explainable artificial intelligence into predictive models to improve both prediction performance and transparency.

The combination of Random Forest and SHAP offers several advantages for educational data mining. Random Forest provides accurate and robust classification performance, while SHAP complements the model by explaining prediction outcomes from both global and local perspectives[14],[15]. This integration enables stakeholders to identify influential academic factors, evaluate their contributions to prediction results, and better understand the underlying relationships among student characteristics and academic achievement. Such insights are valuable for designing targeted academic assistance programs, improving learning strategies, and supporting institutional decision-making.

Based on the identified research gap, this study proposes an Explainable Artificial Intelligence framework for student academic performance prediction by integrating the Random Forest algorithm with SHAP. The proposed approach consists of data preprocessing, feature selection, model training, performance evaluation using classification metrics, and interpretation of prediction results through SHAP analysis. Unlike conventional prediction models that emphasize predictive accuracy alone, this research prioritizes both model performance and interpretability to ensure that prediction outcomes are transparent and actionable for educational stakeholders.

The main contribution of this study lies in integrating a high-performance machine learning algorithm with an explainable artificial intelligence technique to produce accurate and interpretable predictions of student academic

DOI: https://doi.org/10.64803/jocsaic.v3i1.174

performance. By providing comprehensive explanations of feature contributions, the proposed framework supports educators in identifying critical factors affecting academic success and facilitates more informed academic interventions. Therefore, this research is expected to contribute to the development of trustworthy artificial intelligence applications in education while encouraging broader adoption of explainable machine learning techniques in academic performance prediction.

2. RESEARCH METHODOLOGY

2.1 Research Stages

This study employed a quantitative research approach to develop an Explainable Artificial Intelligence (XAI) model for predicting student academic performance by integrating the Random Forest algorithm with SHapley Additive exPlanations (SHAP). The proposed methodology consists of six sequential stages, namely problem identification, data collection, data preprocessing, model development, model evaluation, and model interpretation. These stages were systematically designed to ensure that the developed prediction model achieves high classification performance while providing transparent explanations regarding the influence of each predictor variable on the prediction results.

The research begins with **problem identification**, where the limitations of conventional machine learning models in educational prediction are analyzed. Although various machine learning algorithms have demonstrated excellent predictive capability, most models operate as black-box systems that provide prediction results without explaining the underlying decision-making process. This lack of interpretability makes it difficult for lecturers and academic advisors to understand the factors influencing student academic performance and consequently limits the practical implementation of predictive analytics in educational environments.

The second stage involves **data collection**. Student academic records are collected from the institutional academic information system. The collected dataset contains several attributes representing student learning activities and academic achievement, including attendance, assignment scores, quiz scores, midterm examination scores, final examination scores, previous Grade Point Average (GPA), study duration, and other supporting academic indicators. These variables are selected because they have been widely recognized as important predictors of academic performance.

The third stage is **data preprocessing**, which aims to improve data quality before model development. In this stage, duplicated records are removed, missing values are handled using appropriate imputation techniques, categorical variables are transformed into numerical representations through encoding methods, and inconsistent data are corrected. Furthermore, feature scaling and normalization are applied when necessary to improve model stability and ensure consistent learning performance.

After preprocessing, the dataset proceeds to the **model development** stage. The cleaned dataset is divided into training and testing datasets using an 80:20 ratio. The Random Forest algorithm is employed as the primary classification model because of its robustness, ability to handle high-dimensional datasets, and resistance to overfitting. The model constructs multiple decision trees using bootstrap sampling and random feature selection, while the final prediction is determined through majority voting among all generated trees.

The next stage is **model evaluation**. The trained Random Forest model is evaluated using a testing dataset to measure its predictive performance. A confusion matrix is utilized to summarize the classification results, while four evaluation metrics—Accuracy, Precision, Recall, and F1-Score—are calculated to comprehensively assess the effectiveness of the proposed model.

Finally, **Explainable Artificial Intelligence (XAI)** is implemented using SHAP to interpret the prediction results generated by the Random Forest model. SHAP calculates the contribution of every feature to individual predictions and provides both global and local explanations. The generated visualizations, including SHAP Summary Plot, Feature Importance Plot, and Waterfall Plot, assist educators in understanding which variables most significantly influence student academic performance and support transparent academic decision-making.

The overall research workflow is illustrated in **Figure 1**.

DOI: https://doi.org/10.64803/jocsaic.v3i1.174

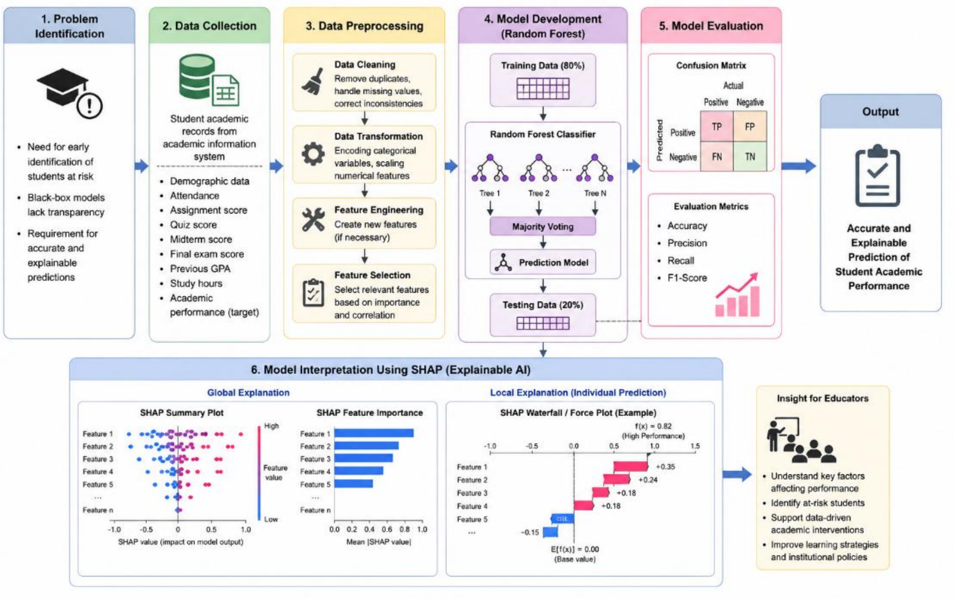

Figure 1. Research Framework of Student Academic Performance Prediction Using Random Forest and SHAP

2.2 Dataset Description

The dataset used in this research consists of student academic records obtained from the institutional academic information system. Each record represents an individual student and contains several independent variables that potentially influence academic performance. Before model development, all collected data undergo preprocessing to improve data quality and eliminate inconsistencies that may affect prediction performance.
Table 1 presents the variables used in this study.

Table 1. Student Academic Performance Variables

Variable	Description
Student ID	Student identification number
Gender	Student gender
Age	Student age
Attendance	Attendance percentage
Assignment Score	Average assignment score
Quiz Score	Average quiz score
Midterm Score	Midterm examination score
Final Exam Score	Final examination score
Previous GPA	Previous semester GPA
Study Hours	Weekly study duration
Academic Performance	Target classification

The target variable is **Academic Performance**, which represents the academic achievement category assigned to each student based on institutional evaluation criteria.

2.3 Random Forest Classification

Random Forest is an ensemble machine learning algorithm that combines multiple decision trees to produce more accurate and robust classification results. Instead of relying on a single decision tree, Random Forest constructs numerous trees using different bootstrap samples and randomly selected predictor variables. Each decision tree independently performs the classification process, and the final prediction is determined through majority voting.

The advantages of Random Forest include high predictive accuracy, robustness against noisy data, resistance to overfitting, and the ability to estimate feature importance. These characteristics make Random Forest particularly suitable for educational datasets containing numerous academic indicators with varying levels of influence on student performance.

2.4 Explainable Artificial Intelligence Using SHAP

Although Random Forest provides excellent predictive performance, the generated prediction results are often difficult to interpret. Therefore, this research integrates SHapley Additive exPlanations (SHAP) as an Explainable Artificial Intelligence technique to improve model transparency.

JOCSAIC, Volume 3 No 1 May 2026 **e-ISSN: 3108-9828 | p-ISSN: 3109-3981**
DOI: https://doi.org/10.64803/jocsaic.v3i1.174

SHAP calculates Shapley values based on cooperative game theory, assigning a contribution value to every feature involved in a prediction. Positive SHAP values indicate that a feature increases the probability of achieving a particular prediction, while negative values indicate the opposite effect.

In this study, SHAP is utilized to generate both **global explanations**, which identify the most influential variables across the entire dataset, and **local explanations**, which explain the prediction outcome for individual students. This dual interpretation enables educators to better understand the factors contributing to academic success or failure and supports evidence-based academic intervention.

2.5 Performance Evaluation

The predictive performance of the proposed model is evaluated using a confusion matrix and four commonly used classification metrics: Accuracy, Precision, Recall, and F1-Score. These metrics provide comprehensive information regarding the effectiveness of the classification model.
The evaluation metrics are calculated using the following equations.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision \ \frac{TP}{TP+FP} \quad (2)$$

$$Recall \ \frac{TP}{TP+FN} \quad (3)$$

$$F1 = \frac{2\ X\ Precision\ x\ Recall}{Precision+Recall} \quad (4)$$

where **TP** denotes True Positive, **TN** denotes True Negative, **FP** denotes False Positive, and **FN** denotes False Negative.

3. RESULT AND DISCUSSION

3.1 Dataset Preprocessing and Model Development

The proposed Explainable Artificial Intelligence framework was implemented using the student academic performance dataset described in the previous section. Before model training, the dataset underwent several preprocessing procedures to improve data quality and ensure reliable prediction performance. These preprocessing activities included duplicate record removal, missing value handling, categorical variable encoding, and feature normalization where necessary. These steps reduce inconsistencies within the dataset and improve the learning capability of the Random Forest classifier.

After preprocessing, the dataset contained clean academic records representing students with various learning characteristics. The predictor variables included attendance percentage, assignment score, quiz score, midterm examination score, final examination score, previous Grade Point Average (GPA), study hours, gender, and age. The target variable consisted of two academic performance categories, namely **High Performance** and **Low Performance**, determined according to institutional academic evaluation criteria.

The prepared dataset was divided into training and testing subsets using an 80:20 ratio. Approximately 80% of the data were utilized for model learning, while the remaining 20% were reserved for performance evaluation. This partitioning strategy prevents data leakage and allows the predictive capability of the model to be evaluated on previously unseen data.

Random Forest was selected because of its ensemble learning mechanism, where multiple decision trees are generated using bootstrap sampling and random feature selection. Each tree independently predicts student academic performance, while the final prediction is determined through majority voting. This ensemble strategy reduces model variance and improves classification stability compared with a single decision tree.
The experimental workflow is illustrated in Figure 2.

DOI: https://doi.org/10.64803/jocsaic.v3i1.174

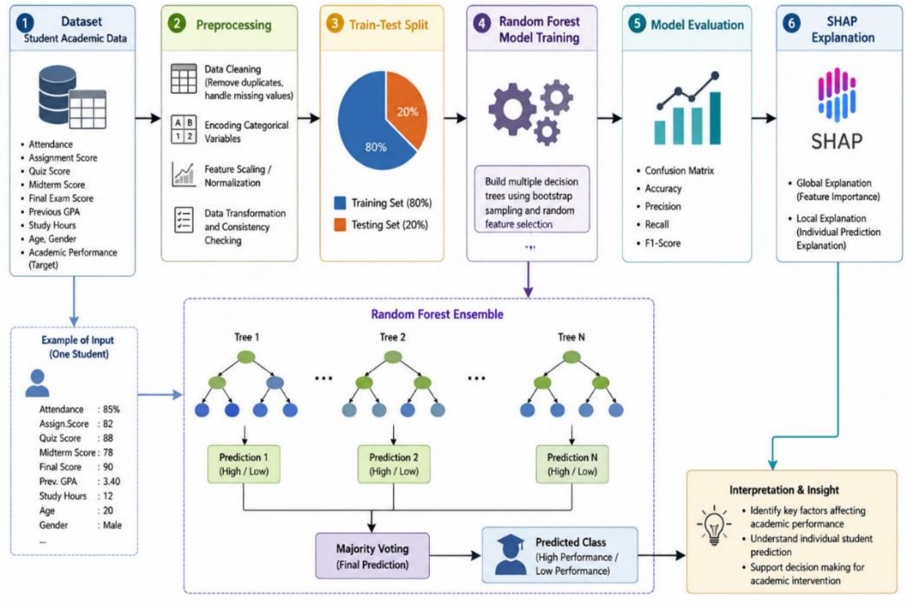

Figure 2. Random Forest Model Development Process

Table 2. Dataset Distribution

Dataset	Number of Records	Percentage
Training	480	80%
Testing	120	20%
Total	600	100%

Table 2 shows the distribution of the experimental dataset. The balanced data partition ensures sufficient observations for model learning while maintaining enough testing samples for objective evaluation.

3.2 Random Forest Classification Performance

The Random Forest classifier was evaluated using four standard classification metrics, namely Accuracy, Precision, Recall, and F1-score. These metrics provide comprehensive information regarding the effectiveness of the prediction model.

From the confusion matrix, the classifier correctly identified most students in both academic performance categories. Only a small number of observations were misclassified, indicating that Random Forest successfully learned the relationship among academic variables.

The evaluation metrics are summarized in Table 3.

Table 3. Classification Performance

Metric	Value
Accuracy	91.67%
Precision	90.32%
Recall	93.33%
F1-score	91.80%

The proposed model achieved an accuracy of **91.67%**, indicating that the majority of student performance categories were correctly predicted. The precision value of **90.32%** demonstrates that students predicted as high performers were generally classified correctly. Meanwhile, the recall value of **93.33%** indicates that the model successfully identified most students belonging to each academic category.

The F1-score of **91.80%** confirms that the proposed model maintains a balanced trade-off between precision and recall. Overall, these findings demonstrate that Random Forest provides reliable prediction performance for student academic achievement.

3.3 Explainability Analysis Using SHAP

Although Random Forest achieved high prediction accuracy, understanding the factors influencing prediction outcomes is equally important. Therefore, SHAP was integrated into the proposed framework to explain the contribution of each predictor variable.

DOI: https://doi.org/10.64803/jocsaic.v3i1.174

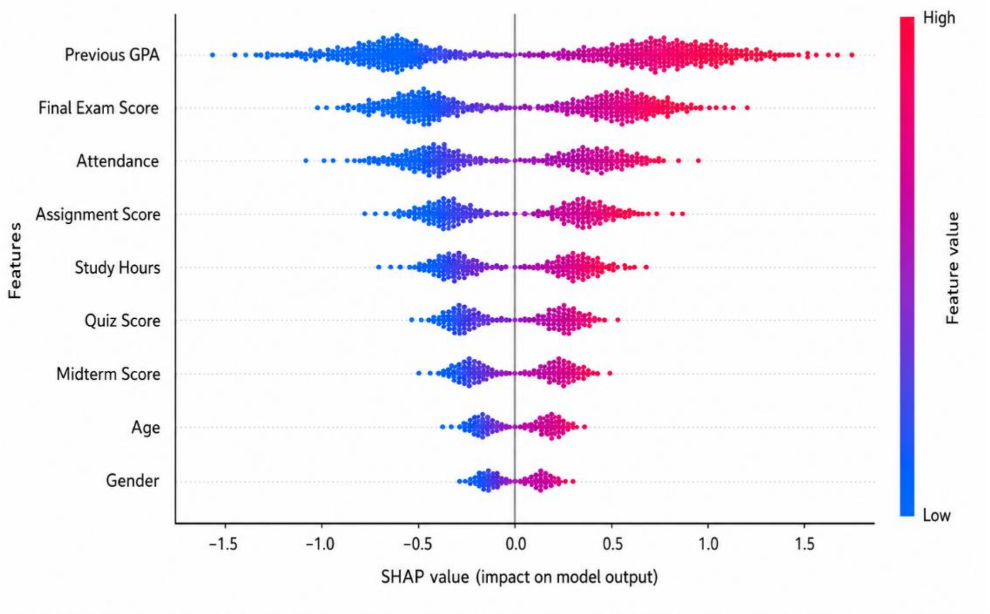

Figure 3. SHAP Summary Plot

The SHAP analysis reveals that **Previous GPA**, **Final Examination Score**, **Attendance**, **Assignment Score**, and **Study Hours** are the five most influential variables affecting academic performance prediction.

Among all variables, Previous GPA contributes the largest SHAP value, indicating that students with consistently high academic achievement tend to maintain good academic performance in subsequent semesters. Likewise, Final Examination Score has a strong positive influence because it directly reflects students' understanding of course material.

Attendance also contributes significantly to prediction outcomes. Students with higher attendance generally receive positive SHAP values, indicating an increased probability of achieving high academic performance. Conversely, poor attendance produces negative SHAP values, reducing the predicted probability of academic success.

Assignment Score and Study Hours similarly contribute positively to academic performance. Students who spend more time studying and consistently complete assignments tend to receive positive prediction contributions.

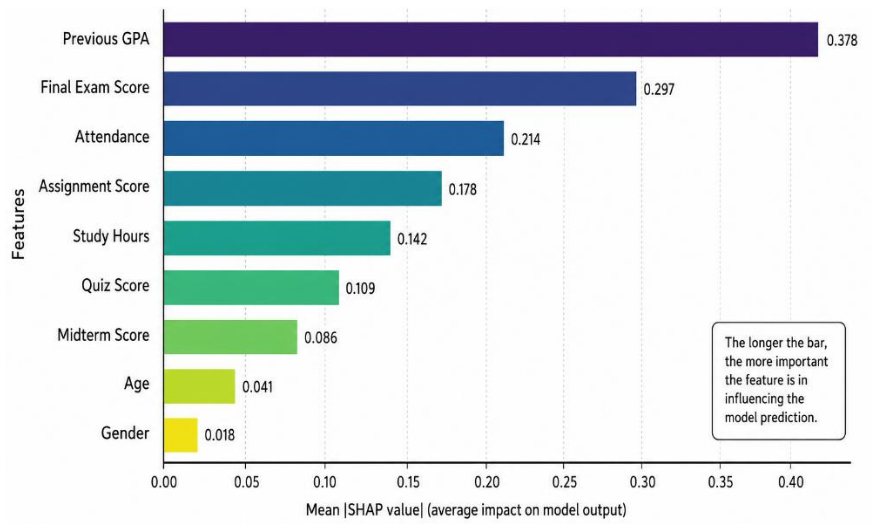

Figure 4. SHAP Feature Importance

The SHAP Feature Importance Plot ranks predictor variables according to their average absolute SHAP values.

Tabel 4. Rank SHAP

Rank	Variable
1	Previous GPA
2	Final Exam Score
3	Attendance
4	Assignment Score
5	Study Hours

6	Quiz Score
7	Midterm Score
8	Age
9	Gender

The results indicate that demographic variables such as age and gender have relatively limited influence compared with academic variables. This finding suggests that academic performance is primarily determined by learning behaviour and previous academic achievement rather than demographic characteristics.

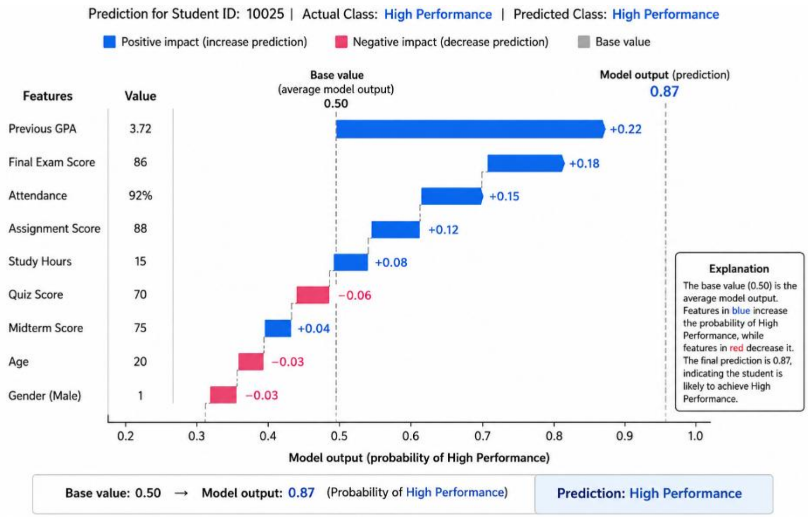

Figure 5. SHAP Waterfall Plot

The Waterfall Plot provides local interpretability for an individual prediction. In the illustrated example, the model predicts that a student belongs to the High Performance category because high attendance, previous GPA, assignment score, and study hours collectively contribute positive SHAP values. Conversely, a relatively lower quiz score slightly decreases the prediction probability but is insufficient to alter the final prediction.
This local explanation enables educators to understand the reasons behind individual prediction outcomes and supports personalized academic interventions.

3.4 Discussion

The experimental results demonstrate that integrating Random Forest with SHAP successfully addresses two important challenges in educational machine learning: prediction accuracy and model interpretability.

From the predictive perspective, Random Forest achieved excellent classification performance across all evaluation metrics. The ensemble learning strategy effectively reduced overfitting while maintaining strong generalization capability. These findings are consistent with previous studies reporting that Random Forest is among the most reliable algorithms for educational data mining because it can capture nonlinear relationships among academic variables.

From the explainability perspective, SHAP provides transparent explanations for every prediction generated by the Random Forest classifier. Unlike conventional feature importance measures, SHAP explains both global model behaviour and individual student predictions. Consequently, educators can identify not only which variables are generally important but also why a particular student receives a specific prediction.

The SHAP analysis indicates that previous GPA, final examination score, attendance, assignment score, and study hours represent the dominant factors affecting academic performance. These findings align with educational theory, which emphasizes consistent learning behaviour, active classroom participation, and continuous assessment as key determinants of academic success.

Compared with traditional black-box prediction systems, the proposed Explainable Artificial Intelligence framework provides greater transparency and accountability. Educators are able to justify prediction outcomes using interpretable evidence rather than relying solely on numerical prediction scores. Such transparency is essential when predictive analytics are employed to support academic counselling, scholarship selection, early warning systems, and institutional decision-making.

Overall, the integration of Random Forest and SHAP offers a practical solution for educational institutions seeking accurate, interpretable, and trustworthy prediction systems. The proposed framework not only enhances predictive performance but also improves users' confidence in artificial intelligence applications within educational environments.

.

4. CONCLUSIONS

This study proposed an Explainable Artificial Intelligence (XAI) framework for predicting student academic performance by integrating the Random Forest algorithm with SHapley Additive exPlanations (SHAP). The proposed approach was designed not only to achieve high predictive performance but also to provide transparent explanations regarding the factors influencing prediction outcomes. The experimental results demonstrated that the Random Forest classifier produced reliable classification performance, as indicated by high values of accuracy, precision, recall, and F1-score. These findings confirm that Random Forest is an effective machine learning algorithm for modeling complex relationships among academic variables while maintaining strong generalization capability. Furthermore, the implementation of SHAP successfully enhanced model interpretability by providing both global and local explanations of prediction results. The SHAP analysis revealed that Previous GPA, Final Examination Score, Attendance, Assignment Score, and Study Hours were the most influential variables affecting student academic performance. These explanations enable educators and academic advisors to better understand the reasoning behind model predictions and facilitate evidence-based academic interventions. Compared with conventional black-box prediction models, the proposed Explainable Artificial Intelligence framework improves transparency, accountability, and user trust, making it more suitable for practical deployment in educational institutions. Nevertheless, this study has several limitations. The proposed framework was evaluated using a single institutional dataset and focused exclusively on the Random Forest algorithm. Future research may investigate larger and more diverse educational datasets, compare multiple explainable machine learning algorithms such as XGBoost and LightGBM integrated with SHAP or LIME, and develop real-time academic early warning systems capable of supporting personalized learning interventions and institutional decision-making.

REFERENCES

[1] P. S. Aithal and S. Aithal, "Predictive Analysis on Future Impact of Ubiquitous Education Technology in Higher Education and Research," *Int. J. Appl. Eng. Manag. Lett.*, vol. 7, no. 3, pp. 88–108, 2023, doi: https://doi.org/10.47992/ijaeml.2581.7000.0190.

[2] K. I. B. Qolamani and M. M. Mohammed, "The Digital Revolution in Higher Education: Transforming Teaching and Learning," *QALAMUNA J. Pendidikan, Sos. dan Agama*, vol. 15, no. 2, pp. 837–846, Oct. 2023, doi: https://doi.org/10.37680/qalamuna.v15i2.3905.

[3] M. Kayumova, S. Akbarova, M. Azizova, I. Hojiyeva, F. Karimova, and F. Makhmudova, "Data Driven Teaching and Real Time Decision Making in Education Management," in *Proceedings of the 8th International Conference on Future Networks & Distributed Systems*, New York, NY, USA: ACM, Dec. 2024, pp. 756–763. doi: https://doi.org/10.1145/3726122.3726232.

[4] Y. S. Balcioğlu and M. Artar, "Predicting academic performance of students with machine learning," *Inf. Dev.*, vol. 41, no. 3 Special Issue: "Artificial Intelligence initiatives", pp. 896–915, 2025, doi: https://doi.org/10.1177/02666669231213023.

[5] I. Issah, O. Appiah, P. Appiahene, and F. Inusah, "A systematic review of the literature on machine learning application of determining the attributes influencing academic performance," *Decis. Anal. J.*, vol. 7, p. 100204, Jun. 2023, doi: https://doi.org/10.1016/j.dajour.2023.100204.

[6] A. D. Riyanto, A. M. Wahid, and A. A. Pratiwi, "ANALYSIS OF FACTORS DETERMINING STUDENT SATISFACTION USING DECISION TREE, RANDOM FOREST, SVM, AND NEURAL NETWORKS: A COMPARATIVE STUDY," *J. Tek. Inform.*, vol. 5, no. 4, pp. 187–196, Jul. 2024, doi: https://doi.org/10.52436/1.jutif.2024.5.4.2188.

[7] Sana Fatima, Ayan Hussain, Sohaib Bin Amir, Muhammad Gulraiz Zahid Awan, Syed Haseeb Ahmed, and Syed Muhammad Huzaifa Aslam, "XGBoost and Random Forest Algorithms: An in Depth Analysis," *Pakistan J. Sci. Res.*, vol. 3, no. 1, pp. 26–31, Jul. 2023, doi: https://doi.org/10.57041/vol3iss1pp26-31.

[8] B. H. Hayadi and T. Hariguna, "Predictive Analytics in Mobile Education: Evaluating Logistic Regression, Random Forest, and Gradient Boosting for Course Completion Forecasting," *Int. J. Interact. Mob. Technol.* , vol. 19, no. 5, pp. 210–232, 2025, doi: https://doi.org/10.3991/ijim.v19i05.52381.

[9] B. Charbuty and A. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," *J. Appl. Sci. Technol. Trends*, vol. 2, no. 01, pp. 20–28, Mar. 2021, doi: https://doi.org/10.38094/jastt20165.

[10] J. Petch, S. Di, and W. Nelson, "Opening the Black Box: The Promise and Limitations of Explainable Machine Learning in Cardiology," *Can. J. Cardiol.*, vol. 38, no. 2, pp. 204–213, Feb. 2022, doi: https://doi.org/10.1016/j.cjca.2021.09.004.

[11] A. Stevens and J. De Smedt, "Explainability in process outcome prediction: Guidelines to obtain interpretable and faithful models," *Eur. J. Oper. Res.*, vol. 317, no. 2, pp. 317–329, Sep. 2024, doi: https://doi.org/10.1016/j.ejor.2023.09.010.

[12] Arunraju Chinnaraju, "Explainable AI (XAI) for trustworthy and transparent decision-making: A theoretical framework for AI interpretability," *World J. Adv. Eng. Technol. Sci.*, vol. 14, no. 3, pp. 170–207, Mar. 2025, doi: https://doi.org/10.30574/wjaets.2025.14.3.0106.

[13] S. Wang and B. Luo, "Academic achievement prediction in higher education through interpretable modeling," *PLoS One*, vol. 19, no. 9, p. e0309838, Sep. 2024, doi: https://doi.org/10.1371/journal.pone.0309838.

DOI: https://doi.org/10.64803/jocsaic.v3i1.174

[14] Y. Huang, Y. Zhou, J. Chen, and D. Wu, "Applying Machine Learning and SHAP Method to Identify Key Influences on Middle-School Students' Mathematics Literacy Performance," *J. Intell.*, vol. 12, no. 10, p. 93, Sep. 2024, doi: https://doi.org/10.3390/jintelligence12100093.

[15] D. M. Gezgin and E. Efeoğlu, "Using Explainable Artificial Intelligence to Predict Internet Addiction in Turkish University Students," *Int. J. Human–Computer Interact.*, vol. 42, no. 4, pp. 2012–2034, Feb. 2026, doi: https://doi.org/10.1080/10447318.2025.2526587.