FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 10 No. 23, December, 2026, pp 420 – 426
DOI: https://doi.org/10.33003/fjs-2026-1023-5393

AI-Powered Student Academic Performance Prediction: A Dual-Model Random Forest Approach for Risk Classification and Score Forecasting

***[1]Aolat Olanike Giwa-Raheem, [1]Abdulbasit Olamide Hussein and [2]Taofik Abiodun Ahmed**

[1]Department of Computer Science, Faculty of Computing, Engineering, and Technology, Al-Hikmah University, Ilorin, Kwara State.
[2]Department of Computer Science, Kwara State College of Arabic and Islamic Legal Studies, Ilorin, Kwara State

*Corresponding authors' email: olanikegiwa@gmail.com

Received: 29 June 2026
Accepted: 26 Oct. 2026
Published: 25 Nov. 2025

ABSTRACT
The increasing availability of student data presents opportunities for educational institutions to identify learners at risk of poor academic outcomes; however, conventional monitoring practices remain largely manual, periodic, and reactive. This study developed and implemented an Artificial Intelligence (AI)-powered system for early academic risk identification and final examination score prediction. Using the UCI Student Performance Dataset, comprising 395 secondary school students and 32 demographic, family, academic, and lifestyle attributes, the study transformed the continuous final grade (G3) into three risk categories: Good, Monitor, and High Risk. A dual-output Random Forest architecture was subsequently developed, comprising a classifier for risk categorisation and a regressor for final score estimation. An 80:20 train-test split, involving 316 training and 79 testing records, was employed with a fixed random seed to support reproducibility. The classifier obtained 86.08% accuracy, a macro F1-score of 0.86, and 0.93 recall for High-Risk students, indicating strong potential for early intervention. The regression model produced a Mean Absolute Error of 1.11 points and an R^2 of 0.83. Feature analysis identified previous-period grades (G1 and G2) as the dominant predictors of final achievement. The models were integrated into a Flask RESTful API and browser-based educator dashboard providing real-time predictions, confidence estimates, and automated explanations. Overall, the findings demonstrate that the proposed dual-model architecture offers an accurate, interpretable, and deployable approach for proactive academic risk management.

Keywords: Artificial Intelligence, Student Performance Prediction, Random Forest, Educational Data Mining, Risk Classification, Machine Learning, Flask API

INTRODUCTION

Today, educational institutions face a growing need to accurately predict student outcomes before final assessments. As universities digitise their operations, they collect massive amounts of student data ranging from demographic details and enrolment statistics to granular records of daily attendance, continuous assessment scores, and socio-economic indicators (Alalawi et al., 2024). Despite this abundance of information, many institutions struggle to translate raw figures into actionable pedagogical insights.

Higher education institutions consider students' academic performance one of the most important issues in providing high-quality education (Wang & Chen, 2025). Understanding the factors that influence student performance is complex, and the academic community has used various effective tools and approaches to address this challenge (Isaeva et al., 2025). In recent years, despite continuous technological progress, a gap remains in the use of machine learning and data mining methods to analyze and improve the accuracy of student performance prediction (Almalawi et al., 2024). Although many researchers have identified factors affecting student performance, the most common predictors remain learning-activity-related variables observed before the final examination (Wang & Yu, 2025).

The traditional administrative models currently in use are largely reactive. They focus on recording what has already occurred rather than anticipating what is likely to happen. Consequently, the identification of academically at-risk students often occurs only after a critical failure event, such as a failed final examination, at which stage remedial intervention is frequently too late to be effective (Carballo-Mendivil et al., 2025). The human and institutional costs of this reactive paradigm are significant, including reduced completion rates, wasted educational resources, and diminished student confidence (Boretti, 2026).

Artificial Intelligence (AI) and Machine Learning (ML) have emerged as transformative tools for addressing this problem. Empirical reviews demonstrate that ML algorithms enable early identification of student performance trajectories, allowing for quicker and more effective interventions (Agyemang et al., 2024). By automating data analysis, such systems allow institutions to shift from reacting to student failure toward actively preventing it.

Among supervised learning algorithms applied to this problem, ensemble methods, and Random Forest in particular, have consistently demonstrated strong performance across diverse institutional contexts and datasets (Santiketa et al., 2024; Turkmenbayev et al., 2025). However, a recurring gap in the literature is the limited attention paid to full-system deployment: most studies report model accuracy without delivering a usable, production-ready interface through which educators can obtain predictions in practice (Hoti et al., 2025).

This study addresses both gaps. It contributes a dual-output Random Forest architecture that simultaneously predicts a categorical risk label and a continuous final score, and it delivers this capability through a complete, deployable web application, a RESTful API backend, and an educator-facing

dashboard rather than a model evaluated only in a research notebook. The remainder of this paper is organized as follows: Section II describes the materials and methods, Section III presents the results and discussion, and Section IV concludes the study.

Review of Related Work

Cortez and Silva (2008) conducted the foundational study on the dataset used in the present work, applying data mining techniques including Decision Trees and regression models to predict the final grades of secondary school students in two Portuguese schools using demographic, social, and school-related attributes. Their best-performing models achieved accuracy levels in the 70–80% range when grade prediction was framed as a binary or coarse-grained classification task. While this study established the value of combining academic, family, and behavioral features for grade prediction, it did not investigate ensemble methods, did not produce a multi-tier risk label distinguishing degrees of academic risk, and did not deliver any form of deployed prediction interface.

Gufroni et al. (2025) compared four supervised learning algorithms: Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), and Extreme Gradient Boosting (XGB) for predicting university admission academic performance using Indonesian SNMPTN pre-admission data combined with GPA records across three train-test schemes. Their Random Forest model achieved the highest accuracy (79%) and precision (0.56), while XGBoost achieved the highest F1-score (0.36) on a 70:30 split, with the resulting best model achieving a 77% success rate on sample validation. A key limitation acknowledged by the authors was severe class imbalance in their four-tier GPA label, which depressed macro-averaged precision and recall across all four algorithms, and the study concluded with model evaluation rather than a deployed system.

Agyemang et al. (2024) conducted an empirical review and practical comparison of five machine learning algorithms for predicting student academic performance, emphasizing careful feature engineering of variables such as attendance rates and continuous assessment scores. Their Random Forest Classifier emerged as the best-performing model, achieving an accuracy of 85.42% and a G-Mean of 0.9243. The study demonstrated that tree-based ensembles can reliably flag at-risk students early enough for low-cost remedial support, but, similar to (Cortez & Silva, 2008; Gufroni et al., 2025), it concluded with offline model evaluation and did not extend to a real-time, web-based prediction interface for educators.

Wang and Yu (2025) proposed a machine learning approach to student performance prediction based on online learning behavioral data, constructing eleven behavioral indicators from the online learning process and statistically filtering out those weakly correlated with student scores. Their resulting logistic regression model, trained on the retained eigenvalue indicators, outperformed five comparative models in prediction ability and confirmed a significant relationship between learning initiative, learning duration, and final performance. While this study contributed a rigorous feature-selection methodology, it relied exclusively on online learning platform data not generally available in conventional classroom settings, and, as with the preceding studies, no deployed prediction system was presented.

Hoti et al. (2025) took a markedly different and highly relevant approach, evaluating user experience rather than predictive accuracy for a web-based student performance prediction platform. Students, professors, and career counsellors tested the deployed platform and reported that it was user-friendly, required minimal technical support, and delivered reliable predictions, with users strongly endorsing its adoption for identifying at-risk students. This study is notable for confirming, through direct user feedback, that deployment and usability, not merely algorithmic accuracy, are essential determinants of real-world adoption. However, the authors explicitly note that the great majority of prior work in this field has focused on prediction algorithm design and evaluation, with comparatively little attention paid to user-centric, deployed evaluation of the kind their study undertakes.

Taken together, these five studies reveal a consistent pattern. Studies that priorities predictive accuracy (Cortez & Silva, 2008; Gufroni et al., 2025; Santiketa et al., 2024) consistently favor Random Forest among the algorithms tested, achieving accuracies in the range of 70–85%, yet none of these studies extends beyond offline model evaluation to a deployed, real-time prediction system. Studies that priorities behavioral feature engineering (Wang & Yu, 2025) improve predictive nuance but depend on data sources unavailable in most institutions. Only the user-evaluation study (Hoti et al., 2025) addresses deployment directly, but it does so for an existing third-party platform rather than proposing or validating a new predictive architecture, and it explicitly identifies the scarcity of user-centric, deployed evaluations as an open problem in the field.

The present study improves on this body of work in three specific ways. First, unlike (Cortez & Silva, 2008; Gufroni et al., 2025; Santiketa et al., 2024), which produce a single prediction output, this study contributes a dual-output Random Forest architecture that simultaneously predicts a three-tier risk category and a continuous final score from a single feature vector, providing richer and more actionable information than a single-output model. Second, the engineered three-tier risk label adopted here (25.3% Good, 41.8% Monitor, 32.9% High Risk) is considerably more balanced than the four-tier GPA label used in (Gufroni et al., 2025), where one class dominated the distribution and depressed classification metrics; this balance contributes to the stronger and more consistent per-class performance reported in Section III. Third, and most significantly, this study closes the deployment gap identified across (Cortez & Silva, 2008; Gufroni et al., 2025; Santiketa et al., 2024), and explicitly flagged in (Hoti et al., 2025): rather than concluding with offline model evaluation, the trained models are deployed through a working RESTful API and a browser-based educator dashboard, delivering the kind of real-time, user-facing prediction capability that (Hoti et al., 2025) confirms is both necessary and currently underexplored in the literature. The remainder of this paper is organized as follows: Section II describes the materials and methods, Section III presents the results and discussion, and Section IV concludes the study.

MATERIALS AND METHODS

This study employs the Random Forest (RF) algorithm in two configurations. Random Forest is an ensemble learning algorithm that combines the outputs of many individual decision trees, producing an output that represents either the class mode (for classification) or the average prediction (for regression) of the trees in the forest (Nachouki et al., 2023). Each tree in the forest is built on a randomly sampled subset of the training data and a random subset of features, which reduces the variance and overfitting that affect individual decision trees (Lam et al., 2020).

Dataset Description

The dataset used in this study is the UCI Student Performance Dataset, originally compiled by Cortez and Silva from two secondary schools in Portugal (Cortez & Silva, 2008). The mathematics subject file was used, containing 395 student records across 33 variables, 32 input features, and one target variable, G3 (the final period grade, scored 0 to 20). The dataset contains no missing values across any of its columns. A summary of the dataset features is presented in Table 1.

Table 1: Summary of Dataset Feature Categories

No.	Category	Example Features	Count
1	Demographic	School, sex, age, address	4
2	Family	Family size, parent status, parental education, jobs, guardian	7
3	School-related	Reason for school, travel time, study time, failures	4
4	Support	School support, family support, paid classes	3
5	Behavioural	Activities, internet, romantic relationship, going out	4
6	Health/social	Family relations, free time, alcohol use, and health	5
7	Academic	Absences, G1, G2	3
8	Target	G3 (final grade, 0–20)	1

Unlike the original SNMPTN admission dataset used in comparable Indonesian university admission research (Gufroni et al., 2025), the UCI dataset reflects continuous in-programme academic and lifestyle indicators rather than pre-admission scores, making it suitable for an in-term early warning use case rather than an admission-stage prediction task.

Data Preprocessing

The data pre-processing stage in this study includes data condition analysis, data cleaning, label encoding, and feature engineering. A visual illustration of this pipeline is presented in Fig. 1.

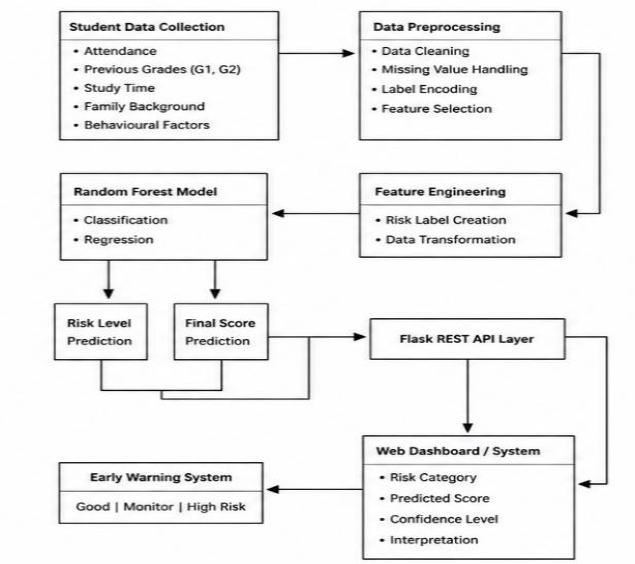


Figure 1: Data preprocessing pipeline of the proposed system

i. Data condition analysis: At this stage, the dataset was checked for missing values, data types, and the distribution of each feature. No missing values were found across any of the 395 records, simplifying the subsequent cleaning stage relative to comparable studies that report substantial missing-value handling requirements (Gufroni et al., 2025).

ii. Label encoding: Seventeen of the thirty-two input features contained categorical text values (e.g., "yes"/"no", "M"/"F", job-type strings) that required conversion into numerical form before being passed to the Random Forest models, which operate exclusively on numerical inputs. Label encoding was applied using the scikit-learn LabelEncoder class (Pedregosa et al., 2011). The specific conversions performed were as follows: binary yes/no fields (schoolsup, famsup, paid, activities, nursery, higher, internet, romantic) were encoded as 1 (yes) and 0 (no); the sex field was encoded as 1 (Male) and 0 (Female); the address field as 1 (Urban) and 0 (Rural); the famsize field as 0 (GT3, greater than 3) and 1 (LE3, 3 or fewer); and the Pstatus field as 1 (Together) and 0 (Apart). Nominal fields with more than two categories; Mjob, Fjob, reason, guardian, and school were encoded as sequential integers assigned alphabetically by the LabelEncoder (e.g., school: GP = 0, MS = 1; guardian: father = 0, mother = 1, other = 2).

iii. Feature engineering — risk label creation: The principal feature engineering step in this study was the derivation of a three-tier risk label from the continuous target variable G3. Table 2 presents the thresholds used and the resulting class distribution.

Table 2: Risk Label Engineering and Class Distribution

Risk Category	G3 Threshold	Count	Percent
Good	G3 ≥ 14	100	25.3%
Monitor	10 ≤ G3 ≤ 13	165	41.8%
High Risk	G3 < 10	130	32.9%

This transformation converts the problem from a pure regression task into a combined classification-and-regression task, enabling the system to output the categorical risk label that is most directly actionable for educators, alongside the continuous score that retains more granular information (Agyemang et al., 2024).

Data Modelling

Following preprocessing, the dataset was partitioned into training and test sets using an 80:20 split ratio (316 training records, 79 test records), with a fixed random seed (random_state = 42) applied via the scikit-learn train_test_split function to ensure full reproducibility (Pedregosa et al., 2011). Unlike the three-scheme comparison (60:40, 70:30, 80:20) adopted in comparable university admission prediction research (Gufroni et al., 2025), this study adopted a single 80:20 split, consistent with the Pareto-optimal ratio commonly applied in machine learning practice when the dataset size is moderate.

Two Random Forest models were trained, each with n_estimators = 100 (100 decision trees per forest) and the same fixed random seed:

i. Random Forest Classifier — trained on the engineered three-tier risk label to predict Good, Monitor, or High Risk.
ii. Random Forest Regressor — trained on the continuous G3 variable to predict the numerical final score.
iii. Both models were implemented using the scikit-learn library in Python (Pedregosa et al., 2011) within a Jupyter Notebook development environment, and serialized using the joblib library for deployment.

Evaluation Metrics

Each model was evaluated using standard metrics appropriate to its task. For the classifier, accuracy, precision, recall, and F1-score were computed using the standard confusion-matrix-based formulae (Gufroni et al., 2025):

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

$$\text{Precision} = TP / (TP + FP) \quad (2)$$

$$\text{Recall} = TP / (TP + FN) \quad (3)$$

$$\text{F1-Score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (4)$$

where TP, TN, FP, and FN denote true positive, true negative, false positive, and false negative counts, respectively, aggregated using macro-averaging across the three risk classes.

For the regressor, Mean Absolute Error (MAE) and the coefficient of determination (R^2) were used:

$$MAE = (1/n)\,\Sigma|y_i - \hat{y}_i| \quad (5)$$

$$R^2 = 1 - (\Sigma\,(y_i - \hat{y}_i)^2 / \Sigma\,(y_i - \bar{y})^2) \quad (6)$$

where y_i is the actual score, $\hat{y}_i$ is the predicted score, $\bar{y}$ is the mean of actual scores, and n is the number of test samples.

System Architecture and Deployment

Beyond model training and evaluation, this study implements a complete three-tier deployment architecture, distinguishing it from related works that conclude at the model evaluation stage (Hoti et al., 2025; Gufroni et al., 2025). The data and model tier consist of the two serialized Random Forest models. The application tier is a RESTful API built with the Python Flask framework (Gowda, 2024), exposing a POST /predict endpoint that accepts a 32-feature JSON payload and returns the predicted risk category, score, per-class confidence, and an automatically generated interpretation string. The presentation tier is a browser-based dashboard implemented in HTML, CSS, and JavaScript, through which an educator enters student data and views the prediction results without requiring any software installation.

RESULTS AND DISCUSSION

Classification Results

The Random Forest Classifier was evaluated on the 79-record test set, achieving an overall accuracy of 86.08%. Table 3 presents the detailed per-class precision, recall, and F1-score values.

Table 3: Classification Report — Random Forest Classifier (Test Set, N = 79)

Class	Precision	Recall	F1-Score	Support
Good	0.96	0.88	0.92	26
High Risk	0.83	0.93	0.88	27
Monitor	0.80	0.77	0.78	26
Macro average	0.86	0.86	0.86	79

The Good class achieved the highest precision (0.96), while the High-Risk class achieved the highest recall (0.93). In the context of an academic early warning system, recall on the High-Risk class is the most operationally significant metric, since a false negative, a genuinely at-risk student misclassified as safe, carries a substantially higher cost than a false positive. The Monitor class returned the lowest F1-score (0.78), consistent with its position as an intermediate category sharing characteristics with both the Good and High-Risk classes, a pattern also observed in comparable multi-class educational prediction studies (Agyemang et al., 2024; Gufroni et al., 2025).

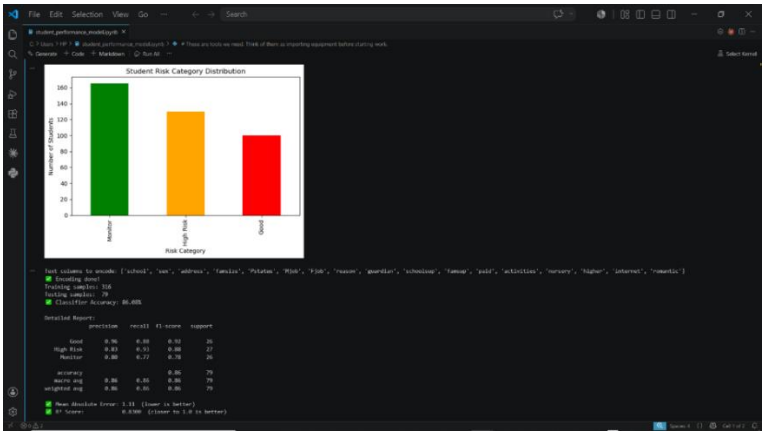


Figure 2: Student risk category distribution

Regression Results

The Random Forest Regressor achieved a Mean Absolute Error of 1.11 grade points (out of 20) and an R-squared value of 0.83 on the test set. This indicates that, on average, predicted scores deviate from actual scores by approximately 5.6%, and that the model accounts for 83% of the variance in final student grades. These results compare favorably with prior regression-based approaches to the same dataset, which typically report accuracy levels in the 70–80% range when grade prediction is framed as a classification task (Cotez & Silva, 2008).

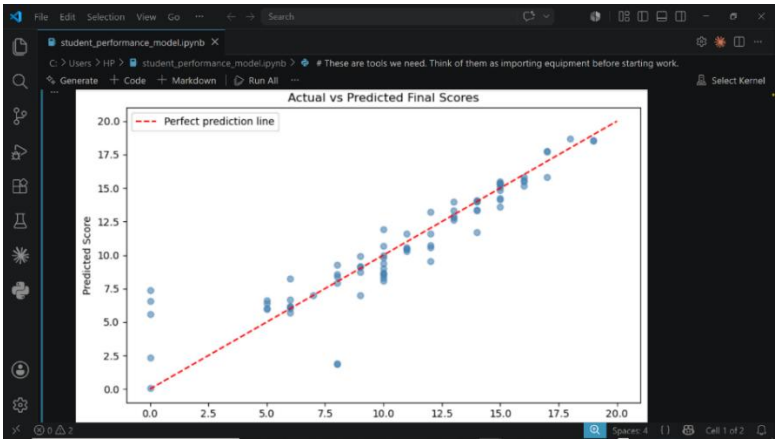


Figure 3: Actual versus predicted final scores on the test set

Feature Importance Analysis

Feature importance analysis was conducted on the trained Random Forest Classifier to identify which input variables most strongly influence the predicted risk category. Table 4 presents the top five most important features.

Table 4: Top 5 Most Important Features — Random Forest Classifier

Rank	Feature	Description
1	G2	Second period grade (0–20)
2	G1	First period grade (0–20)
3	failures	Number of past course failures
4	absences	Number of school absences
5	age	Student age

Consistent with the original dataset study (Cotez & Silva, 2008) and with related work on educational feature importance (Agyemang et al., 2024), the two prior period grades (G2 and G1) emerged as by far the most predictive features. This confirms the intuitive principle that recent academic performance is the strongest available signal for forecasting outcomes, ahead of demographic or lifestyle indicators.

System Deployment and Usability

Beyond model accuracy, the trained models were deployed through a Flask-based RESTful API and a browser-based dashboard. Measured API response times remained consistently below 200 milliseconds under local deployment conditions, confirming the system's suitability for real-time, interactive use during academic advising sessions. Fig. 4 shows the dashboard displaying a sample High Risk prediction, including the predicted score, per-class confidence breakdown, and an automatically generated interpretation statement recommending immediate academic support.

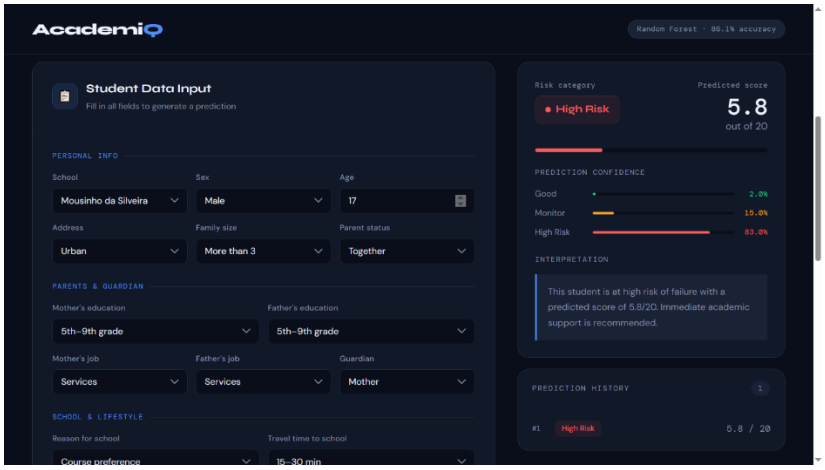


Figure 4: Educator dashboard displaying risk category, predicted score, confidence, and interpretation

Comparison with Related Work

Table 5 situates the present results against comparable studies applying supervised learning to academic performance prediction.

Table 5: Comparison with Related Work

Study	Algorithm(s)	Best Result
Cortez & Silva (2008)	Decision Tree, Regression	~70–80% accuracy
Gufroni et al. (2025)	DT, RF, SVM, XGB	79% accuracy (RF); F1 = 0.36 (XGB)
Agyemang et al. (2024)	Tree-based ensembles	~78–85% accuracy
This study	RF Classifier + Regressor	86.08% accuracy; MAE = 1.11; R^2 = 0.83

The classifier developed in this study achieved a higher accuracy than the comparable university admission study by Gufroni et al. (2025), which reported a best accuracy of 79% using Random Forest on an 80:20 split, and a best F1-score of 0.36 using XGBoost on a 70:30 split. This improvement is attributable in part to the engineered three-tier risk label, which produces a more balanced and discriminative classification target than the four-class, heavily imbalanced GPA-categorization label used in (Gufroni et al., 2025), where the dominant class ("Very Good") constituted the overwhelming majority of records and depressed macro-averaged precision and recall across all four algorithms tested. By contrast, the present study's risk label distribution (25.3% Good, 41.8% Monitor, 32.9% High Risk) is comparatively more balanced, which likely contributed to the stronger and more consistent classification metrics observed across all three classes.

Furthermore, unlike (Cortez & Silva, 2008; Gufroni et al., 2025; Santiketa et al., 2024), which report model evaluation results without a corresponding deployment architecture, the present study additionally delivers a complete, working REST API and dashboard, directly addressing the deployment gap identified across the reviewed literature (Hoti et al., 2025).

Limitations

Several limitations are acknowledged. First, the dataset (n = 395) is drawn from Portuguese secondary schools and may not generalize to other educational contexts without retraining on locally collected data. Second, the Monitor class exhibited comparatively weaker classification performance, suggesting that additional behavioral or engagement features could improve discrimination of borderline cases. Third, the current deployment does not include persistent storage, user authentication, or automated model retraining, each of which represents a direction for future development.

CONCLUSION

This study has demonstrated that a dual-output Random Forest architecture can effectively predict both a categorical academic risk label and a continuous final score from the UCI Student Performance Dataset. The Random Forest Classifier achieved an accuracy of 86.08% and a macro-average F1-score of 0.86, with a recall of 0.93 for the High-Risk class, the most critical class for early intervention purposes. The Random Forest Regressor achieved a Mean Absolute Error of 1.11 points and an R-squared value of 0.83. Feature importance analysis confirmed that prior period grades are the dominant predictors of final academic performance, consistent with the foundational study on this dataset and with subsequent related work.

Beyond model performance, this study contributes a complete, deployable system: A Flask-based RESTful API and a browser-based educator dashboard that together translate the trained models into a practical, real-time decision-support tool. This addresses a recurring gap identified across the reviewed literature, in which academic studies report strong predictive performance without delivering a usable interface for practitioners.

Future work should focus on validating the system against locally collected institutional data, incorporating behavioral and engagement features to improve performance on borderline (Monitor) cases, and extending the architecture with persistent storage, user authentication, and automated periodic retraining to support sustained real-world deployment.

REFERENCES

K. Alalawi, R. Athauda, and R. Chiong, "Harnessing institutional data for actionable pedagogical insight: A review," J. Educ. Data Sci., 2024.

L. Wang and C. J. Chen, "Understanding factors influencing student academic performance," Educ. Technol. Res., 2025.

R. Isaeva, N. Karasartova, K. Dznunusnalieva, K. Mirzoeva, and M. Mokliuk, "Tools and approaches for addressing student performance challenges," Int. J. Educ. Innov., 2025.

A. Almalawi, B. Soh, A. Li, and H. Samra, "Machine learning and data mining for student performance accuracy: A review," Comput. Educ. AI, 2024.

J. Wang and Y. Yu, "Machine learning approach to student performance prediction of online learning," PLOS ONE, vol. 20, no. 1, e0299018, 2025, https://doi.org/10.1371/journal.pone.0299018.

B. Carballo-Mendívil, A. Arellano-González, N. J. Ríos-Vázquez, and M. D. P. Lizardi-Duarte, "Reactive intervention models in higher education," Lat. Am. J. Educ. Tech., 2025.

A. Boretti, "Institutional costs of reactive student support models," Higher Educ. Policy J., 2026.

E. F. Agyemang, J. A. Mensah, O. Ampomah, L. Agyekum, J. Akuoko-Frimpong, A. Quansah, and O. Akinlosotu, "Predicting students' academic performance via machine learning algorithms: An empirical review and practical application," Comput. Eng. Intell. Syst., vol. 15, no. 1, p. 86, 2024, https://doi.org/10.7176/CEIS/15-1-09.

N. Santiketa, S. Chaikhan, U. Ninrutsirikun, and N. Wattanakitrungroj, "Student academic performance prediction using machine learning with various features and scenarios," in Proc. ICSEC 2024, IEEE, pp. 1–6, https://doi.org/10.1109/ICSEC62153.2024.10807584.

A. Turkmenbayev, E. Abdykerimova, S. Nurgozhayev, G. Karabassova, and D. Baigozhanova, "The application of machine learning in predicting student performance in university engineering programs: A rapid review," Front. Educ., vol. 10, art. 1562586, 2025, https://doi.org/10.3389/feduc.2025.1562586.

A. H. Hoti, X. Zenuni, M. Hamiti, and J. Ajdari, "User evaluation of a machine learning-based student performance prediction platform," Organizacija, vol. 58, no. 3, pp. 296–310, 2025, https://doi.org/10.2478/orga-2025-0018.

Elbouknify, I., Berrada, I., Mekouar, L., Iraqi, Y., Bergou, E.H., Belhabib, H., Nail, Y., & Wardi, S. (2025). AI-based identification and support of at-risk students: A case study of the Moroccan education system. *ArXiv, abs/2504.07160*.. https://doi.org/10.48550/arXiv.2504.07160

Kelly, Menezes, A. G., Andre, & Carlos. (2019). *Supervised learning in the context of educational data mining to avoid university students dropout*. in Proc. IEEE ICALT 2019, pp. 207–208, https://doi.org/10.1109/icalt.2019.00068

M. Nachouki, E. A. Mohamed, R. Mehdi, and M. Abou Naaj, "Student course grade prediction using the random forest algorithm: Analysis of predictors' importance," Trends Neurosci. Educ., vol. 33, 2023, https://doi.org/10.1016/j.tine.2023.100214.

K. L. Lam et al., "Use of random forest analysis to quantify the importance of structural characteristics," Food Hydrocoll., vol. 108, p. 106001, 2020, https://doi.org/10.1016/j.foodhyd.2020.106001.

P. Cortez and A. M. G. Silva, "Using data mining to predict secondary school student performance," in Proc. 5th Future Bus. Technol. Conf. (FUBUTEC 2008), pp. 5–12, EUROSIS.

A. I. Gufroni, P. Purwanto, and F. Farikhin, "Academic performance prediction using supervised learning algorithms in university admission," JOIV: Int. J. Inform. Visualization, vol. 9, no. 1, pp. 184–194, 2025, https://doi.org/10.62527/joiv.9.1.2974.

F. Pedregosa et al., "Scikit-learn: Machine learning in Python," J. Mach. Learn. Res., vol. 12, pp. 2825–2830, 2011.

P. Gowda, "Best practices in REST API design for enhanced scalability and security," J. AI Mach. Learn. Des., vol. 2, no. 1, pp. 827–830, 2024, https://doi.org/10.51219/JAIMLD/priyanka-gowda/202.

©2026 This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license viewed via https://creativecommons.org/licenses/by/4.0/ which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is cited appropriately.