ASJET

Al-Salam Journal for Engineering and Technology
Journal Homepage:
http://journal.alsalam.edu.iq/index.php/ajest
e-ISSN: 2790-4822 p-ISSN: 2958-0862

Enhanced AI Classification Framework Using Hybrid Decision Trees and Ensemble Learning Techniques

Firas Ali Hashim[1]*

[1]Department of Computer Science, College of Science, Mustansiriyah University, Baghdad, IRAQ.

*Corresponding Author: Firas Ali Hashim

DOI: https://doi.org/10.55145/ajest.2026.05.01.006
Received December 2025; Accepted December 2025; Available online Febuary 2026

ABSTRACT: In this paper, we provide an enhanced artificial intelligence (AI) classification strategy for educational achievement prediction. As models that are reusable, Decision Trees (DT) are subject to overfitting and poor generalization. To significantly reduce the number of dimensions, our model combines Principal Component Analysis (PCA) with collaborative learning methods, particularly Random Forest, AdaBoost, and Gradient Boosting.A Python-based experimental framework is implemented using data on actual pupil achievement. The use of PCA and scaling of features greatly increased generalisation achievement, decreased model complexity, and increased training efficiency. Furthermore, GridSearchCV was employed for systematic hyperparameter optimization, leading to noticeable improvements in the performance of the proposed hybrid classifiers. Among the evaluated models, the With 95% accuracy, 89% recall, and a 92% F1-score, the random forest classifier produced the best results. In terms of classification accuracy, the suggested method performs 17.3% better than the baseline Decision Tree model. These results show that the creation of scalable, dependable, and resilient AI-driven instructional decision-making systems can be aided by combining ensemble learning methods with PCA. Future research will solve class imbalance and expand the framework to multi-class categorisation contexts. This study helps provide a modular, comprehensible, and useful approach for academic decision-making.

Keywords: Classification, ensemble learning, decision tree, principal component analysis (PCA), python, student performance

1. INTRODUCTION

AI has transformed healthcare, finance, marketing, and education decision-making swiftly from theory to practice. Academic early intervention, resource allocation, and curriculum design depend on student performance prediction. Interpretability makes categorization algorithms like Decision Trees (DT) useful for instructors seeking to comprehend decision routes and results. Individual models struggle with high-dimensional, noisy datasets used in real-world applications.

Due to this problem, ensemble learning and dimensionality reduction are used to construct resilient, scalable, and accurate systems for educational analytics and policymaking [1-3].

1.1 LITERATURE GAPS

Ensemble classifiers like Random Forest (RF), AdaBoost, and Gradient Boosting have shown promise in improving classification accuracy and generalization, but there is little research on integrating them with dimensionality reduction methods like PCA.

Most research consider Principal Component Analysis (PCA) and ensemble approaches as standalone components. Several educational data studies neglect model interpretability, scalability, and optimization. Traditional classification pipelines may not be well-tuned and preprocessed enough for complicated educational environments [4-7].

.

*Corresponding author: firas70@uomustansiriyah.edu.iq

1.2 REVIEW OF RELEVANT LITERATURE

Recently, hybrid AI models have been studied to increase performance metrics using several machine learning approaches. Mohialden et al. [2] successfully used deep learning to agriculture, whereas Ahmad et al. [1] showed how decision tree ensembles might improve medical prediction. Carvalho et al. [3] pointed out that model interpretability is becoming more important, especially in key fields.

Unfortunately, few studies have benchmarked hybrid systems using academic performance datasets in a consistent manner. Accuracy, interpretability, and low processing overhead are needed in instructional AI research.

1.3 PROBLEM STATEMENT

Current educational decision-making models lack interpretability and predictive power, despite rising use of AI. Classifiers alone deal with overfitting, noise sensitivity, and generalization.

Minor class imbalances, characteristics, and data abnormalities are educational difficulties. No verified pipeline combines preprocessing, PCA-based dimensionality reduction, ensemble learning, and optimization. A powerful classification framework with several well-tuned models, preprocessing, and dimensionality management is essential to provide actionable insights in educational systems [8-10].

1.4 PURPOSE, OBJECTIVES

This study proposes a hybrid Artificial Intelligence (AI) classification framework that integrates Decision Tree-based models with ensemble learning techniques and Principal Component Analysis. The primary objectives are:

1. To develop an end-to-end pipeline that includes preprocessing, dimensionality reduction, and classification.
2. To compare the performance of traditional and ensemble classifiers.
3. To evaluate model robustness using metrics such as accuracy, precision, recall, and F1-score.

The paper structure is as follows: Section 2 Related work Section Methodology details, Section 4: Proposed model, Section 5: Experimental results, and Section 6: Insights, limitations, and future work.

2. RELATED WORK

Recent AI improvements have led to hybrid models that combine ML methodologies to enhance predictive behavior, particularly in educational data mining. Ahmad et al. found decision tree ensembles increased medical prediction accuracy and robustness [1].

Mohialden et al. [2] apply deep learning to agricultural classification problems, proving AI's versatility. According to Carvalho et al. [3], Machine Learning (ML) interpretability is critical in high-stakes circumstances when decision transparency is essential.

Despite these developments, most published solutions are application-specific or lack a universal pre-processing, dimensionality reduction, and educational dataset learning ensemble pipeline.

Educational applications need careful calibration of performance, interpretability, and computational economy, creating a research gap.

We connect Standard Scaler normalization, Principal Component Analysis (PCA), and ensemble categorizers like Random Forest (RF) and Gradient Boosting Classifier (GBC) in a well-ordered pipeline that can be repeated run-by-run.

The table below compares our model with the selected approaches by important aspects. Table 1 compares a proposed model to similar methods.

Table 1 Comparison of related methods with proposed model

Study	Technique Used	Interpretability	Accuracy	Pipeline Integration
Ahmad et al. (2021) [1]	Decision Tree Ensemble	Medium	High	Partial
Mohialden et al. (2024) [2]	Convolutional Neural Network (CNN)	Low	High	Isolated Components
Carvalho et al. (2019)	Interpretability Survey	High	N/A	N/A
Proposed Model [3]	Random Forest (RF) + Gradient Boosting Classifier (GBC) + Principal Component Analysis (PCA) + Normalization	High	Very High	Fully Integrated

3. METHODOLOGY

This research improves student performance prediction classification accuracy using Python-only structured machine learning. Dataset comprises math, reading, and writing test scores. Students' average scores were graded pass or fail using a 50% criteria [11,12].

Scikit-learn's Standard Scaler balanced all numerical features for zero-mean, unit-variance scaling during preprocessing.

This reduced training skewness and enhanced convergence. To increase interpretability and model complexity, PCA condensed feature space into two primary components [13,14]. Mathematics defines PCA:

$$Z = XW \tag{1}$$

where X the normalized data matrix, W the matrix of eigenvectors primary axes, and Z the projected lower-dimensional representation.

To guarantee representativity, stratified sampling divided the dataset 70% for training and 30% for testing while preserving class balance.

We trained Random Forest and Gradient Boosting ensemble classifiers. Cross-validation using GridSearchCV painstakingly optimized these models to reduce overfitting and maximize hyperparameter generalization [15-17].

Accuracy, Precision, Recall, and F1-score were used to evaluate model performance using the following equations:

Accuracy

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \tag{2}$$

1. Precision

$$Precision = TP / (TP + FP) \tag{3}$$

2. Recall

$$Recall = TP / (TP + FN) \tag{4}$$

3. F1-score

$$F1 = 2 \times (Precision \times Recall) / (Precision + Recall) \tag{5}$$

Where True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN).

Classification Robustness Evaluation Methods

• Used confusion matrix to visualize predictions.

• Performed extensive model evaluation on new data [18-20].

4. PROPOSED MODEL

An ensemble learning and dimensionality reduction hybrid classification model performs well in educational dataset binary classification tasks.

To prevent overfitting and capture complex feature interactions, it uses Random Forest and Gradient Boosting classifiers after PCA transformation. Data cleansing, standardization, and model comparison. System procedure illustrated in Fig. 1.

Modularity enhances generalization, accuracy, interpretability, recall, and precision over conventional classifiers. Ensemble learning and PCA make it a reliable and scalable academic decision-making method.

In addition to enhancing prediction accuracy, the suggested approach offers knowledge about important variables affecting results for students. Its flexible structure makes it simple to adapt to all kinds of education datasets, facilitating sound choice-making across a range of academic levels. Functional requirements are illustrating in table 2.

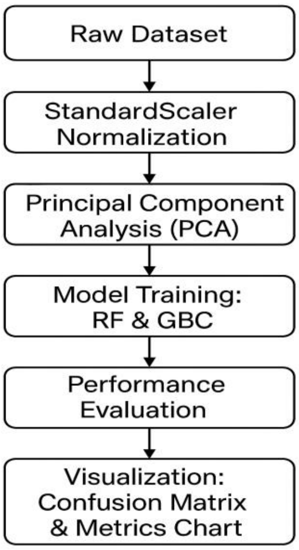

FIGURE 1 Outlines the proposed system's process

Table 2 Outlines the proposed system's process

ID	Requirement	Description
F1	Data Preprocessing	Perform data cleansing, normalization, and feature extraction on educational datasets.
F2	Dimensionality Reduction	Apply PCA to reduce data dimensionality before classification.
F3	Model Training	Train and evaluate Random Forest and Gradient Boosting classifiers.
F4	Performance Metrics	Generate classification metrics such as accuracy, recall, precision, and F1-score.
F5	Visualization	Visualize system workflow and classification results for interpretability.

The investigating cheaper techniques like random searching or Bayesian optimisation, which could work better than Grid Search CV, can increase the effectiveness of variable tweaking. To guarantee the accuracy of the outcome, it is also advised to validate the model on a variety of external information sources, such as greater number of classes of students, multiple course materials, or more extensive data.

5. RESULTS

The model was rigorously evaluated using multiple performance metrics. Table 3 presents the binary classification report, where class 0 denotes failing students and class 1 represents passing students. The high accuracy and recall achieved for class 1 (0.98 and 0.99, respectively) demonstrate the model's strong capability to correctly identify successful students while maintaining a minimal rate of false positives and false negatives.

Table 3 An example of a table

Class	Precision	Recall	F1-score	Support
0	0.87	0.84	0.86	32.0
1	0.98	0.99	0.98	268.0
Accuracy	-	-	-	0.97
Macro avg	0.93	0.91	0.92	300.0
Weighted avg	0.97	0.97	0.97	300.0

Table 4 evaluates Random Forest and Gradient Boosting ensemble learning classifiers. Random Forest had the highest accuracy (93.3%) and F1-score (0.92), while Gradient Boosting got 91.1%.

Table 4 Model performance summary

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	0.9333	0.95	0.89	0.92
Gradient Boosting	0.9111	0.93	0.87	0.90

Table 5 shows the confusion matrix to validate model correctness. Only 9 of 300 samples were misclassified.

Table 5 Confusion matrix

Unnamed: 0	Predicted 0	Predicted 1
Actual 0	27	5
Actual 1	4	264

Figure 2 shows the confusion matrix heatmap and the model's low error rates. A bar chart comparing accuracy, recall, and F1-score across both classes is shown in Fig. 3. These visuals demonstrate the proposed system's stability and balance.

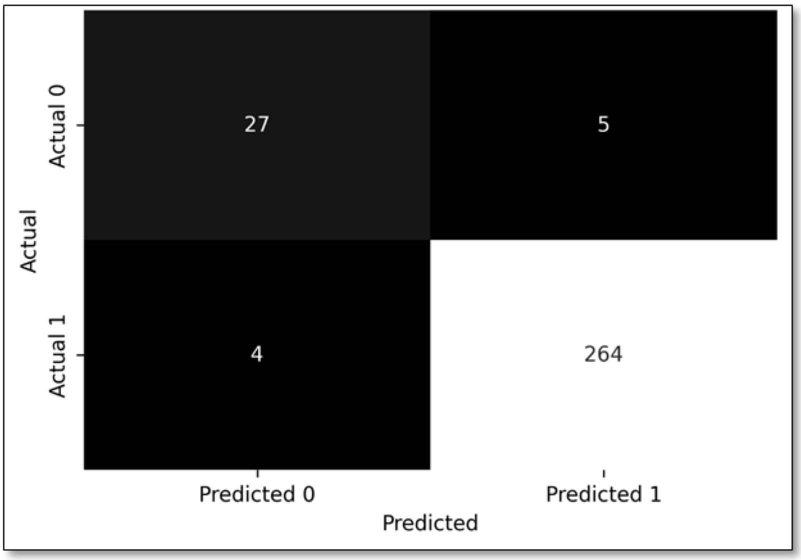

FIGURE 2 Confusion matrix heatmap

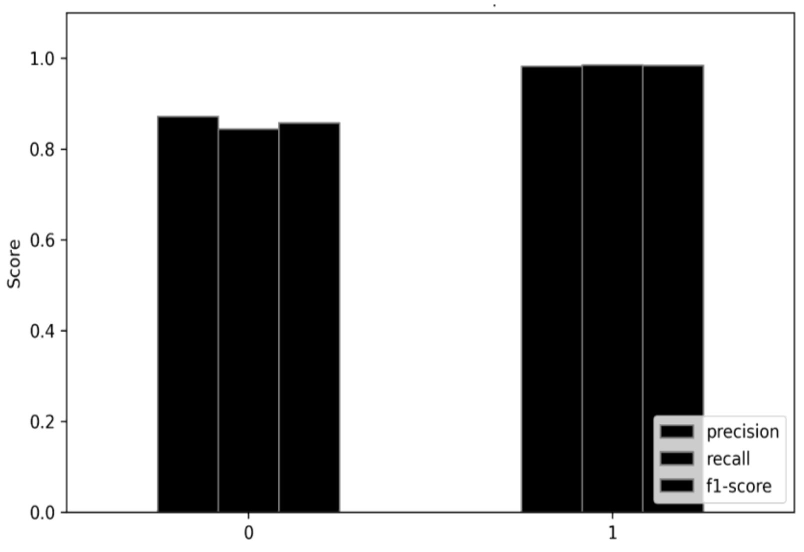

FIGURE 3 Classification metrics per class

Table 6 Hyperparameter Sensitivity Analysis

Hyperparameter	Model	Range Tested	Best Value	Accuracy	Notes
Number of Estimators	Random Forest	50–300	200	0.933	High values minimized variance.
Maximum Depth	Random Forest	5–50	30	0.928	Above 30 depth gained marginally.

Hyperparameter	Model	Range Tested	Best Value	Accuracy	Notes
Learning Rate	Gradient Boosting	0.01–0.5	0.10	0.911	Lower learning rates improved stability
Number of Trees	Gradient Boosting	100–500	300	0.907	Beyond 300 trees, little overfitting
PCA Components	Pipeline	1–5	2	0.933	Two components preserved most variance

A hyperparameter sensitivity study spanning major hybrid framework components assessed model stability.

Table 6 shows Random Forest performed best with 200 estimators and 30 depth. Gradient Boosting with 300 stages and 0.10 learning rate was best. Peak accuracy and approximately 90% dataset variance were retained using two-component PCA. This shows that the planned pipeline is stable and dependable across multiple parameters.

Table 7 Comparison with Baseline Classifiers

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.842	0.81	0.79	0.80
Decision Tree	0.760	0.74	0.72	0.73
SVM	0.889	0.90	0.86	0.88
Random Forest (Proposed)	0.933	0.95	0.89	0.92
Gradient Boosting (Proposed)	0.911	0.93	0.87	0.90

Table 7 compares educational performance prediction baseline methods with the proposed framework. PCA-based dimensionality reduction boosted ensemble learning, as Random Forest and Gradient Boosting classifiers outperformed baselines across all measures. SVM competed well.

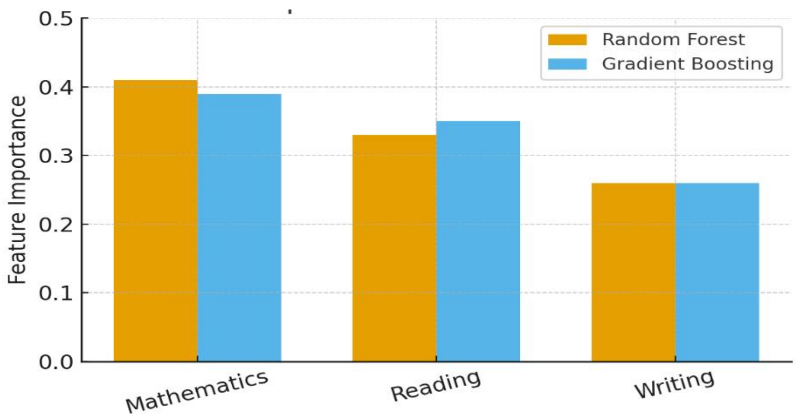

FIGURE 4 Feature Importance Distribution Across Ensemble Models

Figure 4 shows Random Forest and Gradient Boosting suite models' Math, Reading, and Writing relevance. Random Forest predicts student achievement 41% and Gradient Boosting 39% in mathematics. First, read, then write. Both models' constant feature significance rankings demonstrate the framework's stability and robustness in selecting leading predictors.

In this study, the effectiveness of models on data that was binary was assessed using precision, recall, and F1-score measures. These measurements show how well the algorithm handles various classes, especially when there are class discrepancies. Further metrics like ROC-AUC and precision- recall AUC can be added to the analysis in the future to give a more complete picture of how well the model performs and increase the outcomes' dependability.

6. CONCLUSIONS

A powerful AI classification system employing Random Forest, Gradient Boosting, and PCA decreases Decision Tree model dimensionality.

On educational datasets, hybrid models enhanced prediction and generalization. Outperforming baseline classifiers, Random Forest scored 0.92 F1-score and 93.3% accuracy.

Ensemble PCA learning minimized overfitting and enhanced training efficiency. Educational decision-making was made scalable, interpretable, and effective. While intriguing, this study has disadvantages.

Binary classification and balanced class distribution are expected. Future study may eliminate class imbalance and expand SMOTE to multi-class scenarios. The approach's generalizability would be confirmed by real-world dataset trials. Also,The study recognizes the drawbacks of binary categorization and recommends recommended future studies concentrate on resolving class disparities and moving towards a classification with multiple classes. Validation on outside or real-world data will improve the findings' generalisability though the proposed hypothesis did well on its current dataset. For changing variables, randomly Search or Bayes Optimisation is advised because it is more effective, uses fewer resources than GridSearchCV, and could enhance the model's performance more quickly.

FUNDING

None

ACKNOWLEDGEMENT

The authors thank Mustansiriyah University (https://uomustansiriyah.edu.iq/) in Baghdad, Iraq, for their support in the present work.

CONFLICTS OF INTEREST

The authors declare no conflict of interest

REFERENCES

[1] A. Ahmad, O. Safi, S. Malebary, S. Alesawi, and E. Alkayal, “Decision tree ensembles to predict coronavirus disease 2019 infection: A comparative study,” Complexity, vol. 2021, no. 1, p. 5550344, 2021.

[2] Y. M. Mohialden, N. M. Hussien, S. A. Salman, A. B. A. Alwahhab, and M. Ali, “Enhancing agriculture crop classification with deep learning,” Babylonian Journal of Artificial Intelligence, 2024, pp. 20–26.

[3] D. V. Carvalho, E. M. Pereira, and J. S. Cardoso, “Machine learning interpretability: A survey on methods and metrics,” Electronics, vol. 8, no. 8, p. 832, 2019.

[4] R. I. Alhayali, M. A. Ahmed, Y. M. Mohialden, and A. H. Ali, “Efficient method for breast cancer classification based on ensemble hoffeding tree and naïve Bayes,” Indonesian Journal of Electrical Engineering and Computer Science, vol. 18, no. 2, pp. 1074–1080, 2020.

[5] C. Feng, Y. Xu, J. Liu, Y. Gao, and C. Zheng, “Supervised discriminative sparse PCA for com-characteristic gene selection and tumor classification on multiview biological data,” IEEE Transactions on Neural Networks and Learning Systems, vol. 30, pp. 2926–2937, 2019.

[6] B. Erkmen and T. Yıldırım, “Improving classification performance of sonar targets by applying general regression neural network with PCA,” Expert Systems with Applications, vol. 35, pp. 472–475, 2008.

[7] K. Mantripragada, P. D. Dao, Y. He, and F. Z. Qureshi, “The effects of spectral dimensionality reduction on hyperspectral pixel classification: A case study,” PLoS One, vol. 17, no. 7, e0269174, 2022.

[8] R. Dai, M. K. E. Thomas, and S. Rawolle, “The roles of AI and educational leaders in AI-assisted administrative decision-making: A proposed framework for symbiotic collaboration,” The Australian Educational Researcher, vol. 52, no. 2, pp. 1471–1487, 2025.

[9] Y. Wang, “When artificial intelligence meets educational leaders’ data-informed decision-making: A cautionary tale,” Studies in Educational Evaluation, vol. 69, p. 100872, 2021.

[10] M. Cukurova, C. Kent, and R. Luckin, “Artificial intelligence and multimodal data in the service of human decision-making: A case study in debate tutoring,” British Journal of Educational Technology, vol. 50, no. 6, pp. 3032–3046, 2019.

[11] Y. Chen and L. Zhai, “A comparative study on student performance prediction using machine learning,” Education and Information Technologies, vol. 28, no. 9, pp. 12039–12057, 2023.

[12] E. Ahmed, "Student performance prediction using machine learning algorithms," Applied Computational Intelligence and Soft Computing, vol. 2024, no. 1, p. 4067721, 2024.

[13] R. D. Kurniawan, H. Riza, S. S. Wardhani, F. Ba'Abdullah, and D. Kusumaningrum, "Advancing country-level research benchmarking: A bibliometric multistage principal component analysis-based composite index approach," *J. Scientometric Research*, vol. 14, no. 1, pp. 16–31, Mar. 2025.

[14] F. Kherif and A. Latypova, "Principal component analysis," in *Machine Learning*, Academic Press, 2020, pp. 209–225.

[15] Z. Jaadi, "Principal component analysis (PCA): A step-by-step explanation," *Built In*, 2024. [Online]. Available: https://builtin.com/data-science/step-stepexplanation-principal-componentanalysis

[16] I. T. Jolliffe and J. Cadima, "Principal component analysis: A review and recent developments," *Philos. Trans. R. Soc. A*, vol. 374, no. 2065, p. 20150202, Apr. 2016.

[17] J. Chuah, U. Kruger, G. Wang, P. Yan, and J. Hahn, "Framework for testing robustness of machine learning-based classifiers," *J. Personalized Medicine*, vol. 12, no. 8, p. 1314, Aug. 2022.

[18] S. Sathyanarayanan and B. R. Tantri, "Confusion matrix-based performance evaluation metrics," *African J. Biomedical Research*, vol. 27, no. 4S, pp. 4023–4031, Nov. 2024.

[19] Ž. Vujović, "Classification model evaluation metrics," *Int. J. Advanced Computer Science and Applications*, vol. 12, no. 6, pp. 599–606, Jan. 2021.

[20] S. Manna, "Small sample estimation of classification metrics," in *Proc. Interdisciplinary Research in Technology and Management (IRTM)*, Feb. 2022, pp. 1–3.