



Comparative Analysis of Decision Tree and Random Forest Models for Student Exam Score Classification

Pocut Naura Nisrina^{1*}, Muhammad Rizqi Antara², Chelsy Intami³,
Asmara Azkia Nurazizah⁴, Mujahidah Fathul Islam⁵, Sania Rivka Madina⁶,
Ratu Fatimah Zahro⁷, Isna Jami'atul Khoeriyah⁸, Ahmad Afif⁹, Nur Ilfi Aisah¹⁰

^{1,2,3}Department of Information System, Faculty of Science and Technology,
Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia

⁴Department of Ushul Fiqh, Faculty of Sharia and Islamic Studies,
Al Qasimia University, United Arab Emirates

⁵Department of History, Faculty of Social Sciences and Humanities,
Sakarya University, Türkiye

⁶Department of Islamic Studies, Faculty of Humanities and Social Sciences,
Ibnu Tofail University, Morocco

⁷Department of Dirasat Islamiyah, Faculty of Syari'ah Islamiyah,
University of Al Azhar, Egypt

⁸Department of Ushuluddin, Faculty of Tafseer and Quran Science,
International Islamic University of Islamabad, Pakistan

⁹Magister of Arts, University of Al-Azhar, Egypt

¹⁰Bachelor of Media and Communication (Hons), School of Education and Human Sciences,
Albukhary International University, Malaysia

E-Mail: ¹12350321126@students.uin-suska.ac.id, ²12350313161@students.uin-suska.ac.id,
³12350320442@students.uin-suska.ac.id, ⁴q23124983@alqasimia.ac.ae,
⁵mujahidah.islam@ogr.sakarya.edu.tr, ⁶saniarivka.madina@uit.ac.ma,
⁷ratufatimahzahro6112@gmail.com, ⁸012112bsudtqsf25@student.iiu.edu.pk,
⁹ahmadafifa522@gmail.com, ¹⁰nurilfiaisah@gmail.com

Received Dec 30th 2025; Revised Jul 18st 2026; Accepted Jul 26th 2026; Available Online Aug 01st 2026

Corresponding Author: Pocut Naura Nisrina

Copyright © 2026 by Authors, Published by Institute of Research and Publication Indonesia (IRPI)

Abstract

Advances in Educational Data Mining (EDM) have encouraged the use of machine learning algorithms to predict students' academic performance. Although Decision Tree and Random Forest have been widely investigated, most previous studies compare them with multiple classifiers rather than conducting a focused evaluation under a unified framework. This study addresses this gap by systematically comparing the performance of Decision Tree and Random Forest in classifying student exam scores using the same preprocessing, cross-validation, and evaluation procedures. The dataset, obtained from Kaggle, consists of six attributes: student_id, hours studied, sleep hours, attendance percent, previous scores, and exam scores. The research methodology includes data collection, preprocessing, K-Fold Cross Validation, model training, and performance evaluation using a confusion matrix with accuracy, precision, and recall metrics. Experimental results show that Random Forest consistently outperforms Decision Tree, achieving a mean accuracy of 86.50%, compared with 81.50% for Decision Tree. In addition, Random Forest demonstrates greater stability and stronger generalization capability across different validation folds. These findings suggest that Random Forest is a more effective and reliable predictive model for supporting educational decision-making, particularly in identifying students at risk of poor academic performance and enabling timely academic interventions.

Keywords: Decision Tree, Educational Data Mining, Exam Score Classification, Random Forest, Student Performance

1. INTRODUCTION

The development of information technology has driven the widespread use of Educational Data Mining (EDM) to uncover learning patterns, predict academic performance, and support data-driven educational decision-making. EDM enables in-depth analysis of student learning behavior using various machine learning techniques that transform complex data into meaningful insights [1]. Numerous studies highlight that

algorithms such as Decision Tree, Random Forest, and Support Vector Machine (SVM) are among the most frequently used and proven effective for academic classification tasks [2].

Despite the increasing availability of educational data, many educational institutions still struggle to accurately identify the factors that influence student exam performance. Traditional evaluation methods often fail to capture hidden patterns in academic data, leading to difficulties in the early detection of low-performing students and in planning effective academic interventions. These challenges highlight the need for predictive models that are not only accurate but also interpretable to support informed decision-making by educators and administrators.

To address this issue, machine learning algorithms offer a promising solution. Linear SVM, for example, has demonstrated strong performance in predicting learning outcomes on both categorical and binary datasets, especially in school-based educational settings [3]. Other research also emphasizes that academic, psychological, and socioeconomic factors significantly influence the accuracy of model predictions and should be considered in the analysis [4]. Comparative studies involving KNN, SVM, and Decision Tree further show that these models can identify influential variables related to student performance, including attendance, family background, and socioeconomic conditions [5]. While these studies confirm the predictive value of behavioral and socioeconomic features, most report performance on a single train-test split without cross-validation, which limits the reliability of their accuracy estimates and makes it difficult to assess how well the models generalize to unseen data.

Decision Trees and Random Forests remain popular in EDM due to their interpretability and robust predictive performance. Random Forest frequently outperforms Decision Tree in various classification studies due to its ensemble nature and resistance to overfitting [6]. Several studies show that attributes such as attendance, assignment scores, and previous-semester grades are strong predictors in Decision Tree-based models [7]. Random Forest has also demonstrated excellent predictive power for course-grade forecasting, achieving up to 90.33% accuracy in prior research [8]. Recent comparative analyses indicate that Random Forest consistently surpasses Decision Tree in classification tasks [9], while SVM, particularly with a linear kernel, achieves the highest accuracy and even outperforms Random Forest in academic grade prediction [10]. Similar findings have been reported in studies employing diverse algorithms such as Logistic Regression, SVM, Random Forest, and XGBoost, showing that support-vector-based models often deliver optimal performance in predicting student grades [11], [12], [13]. These mixed findings indicate that the relative performance of Decision Tree, Random Forest, and SVM is dataset-dependent, and they reveal a limitation shared by much of this literature: because these studies typically benchmark four or more algorithms simultaneously, they rarely provide an in-depth, head-to-head analysis of Decision Tree against Random Forest specifically, leaving the trade-off between these two frequently cited top performers comparatively underexplored.

Other international studies support these conclusions. Dervenis et al. [14] showed that combining learning behavior factors with supervised learning algorithms enhances prediction accuracy. Predictive models of student failure using algorithms such as SVM, Decision Tree, and Naïve Bayes have also been successfully implemented in higher-education contexts [15]. Recent comparative research further highlights that algorithms such as SVM, Random Forest, and Gradient Boosted Trees exhibit strong performance in forecasting final exam results [16], while other work confirms that SVM and Random Forest remain among the most stable algorithms for academic performance prediction across educational contexts [17], [18]. However, these studies are largely confined to a single institutional dataset or a specific educational level, which constrains the generalizability of their conclusions. Few explicitly examine the stability of model performance across repeated folds or resampling, a gap that a cross-validated comparison can help close.

Based on these findings, Decision Tree and Random Forest are appropriate algorithms to compare in this study, as each provides unique strengths in interpretability, stability, and predictive accuracy [19]. Taken together, the reviewed literature shows consistent but sometimes conflicting evidence regarding the relative performance of Decision Tree and Random Forest, and few studies isolate these two algorithms under a single, consistently validated framework specifically designed for exam score classification. Therefore, this research aims to analyze the performance of these algorithms for predicting student exam scores, with the goal of producing an accurate predictive model and strengthening academic decision-making. The novelty of this study lies in combining these algorithms within a unified evaluation framework tailored specifically for exam score classification, enabling a comprehensive, focused comparison that may guide institutions in selecting the most effective predictive approach.

2. MATERIAL AND METHOD

The purpose of the elements in the diagram is to structure and organize the data, to improve model accuracy and performance, and to minimize bias and error in the analysis of results. Research Methodology can see Figure 1

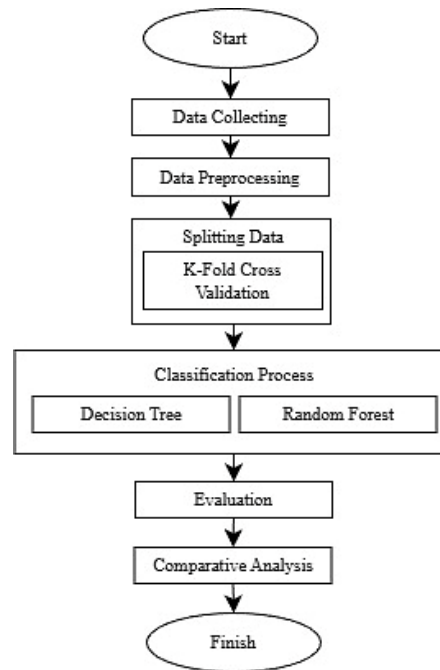


Figure 1. Research Methodology

The research method illustrated in the diagram comprises several interconnected stages from start to finish. The process begins with the identification of objectives and data requirements. The first stage is data collection, where relevant data is gathered from the predetermined sources. Once the data is obtained, data preprocessing is conducted to clean and prepare the dataset, including handling missing values, normalizing data, encoding categorical attributes, and removing inconsistencies that may affect model performance.

After preprocessing, the dataset is divided using K-Fold Cross Validation, which splits the data into several folds to ensure more reliable and stable model evaluation while minimizing the risk of overfitting. The next stage is the classification process, in which two machine learning algorithms, Decision Tree and Random Forest, are trained and tested on the prepared data.

Once the classification results are generated, the evaluation stage assesses each algorithm's performance using metrics such as accuracy, precision, recall, and F1-score. The findings from the evaluation are then compared in the comparative analysis stage to determine which algorithm performs better in terms of predictive accuracy and model stability. The research process concludes with the final stage, in which overall insights and conclusions are drawn from the results obtained.

2.1. Data Collecting

The data in this study is an analysis of a student exam score dataset obtained from Kaggle. This dataset consists of 6 attributes, including numerical and categorical data: Student ID, Hours Studied, Sleep Hours, Attendance Percent, Previous Scores, and Exam Score.

2.2. Data Preprocessing

After data has been collected, the next step is to prepare it for machine learning analysis. In this preprocessing stage, several common procedures are performed, such as filling in missing values, removing duplicate entries, checking for inconsistencies, cleaning the data, and correcting any errors. In this study, the dataset was already well-structured and contained no missing values, so no additional cleaning was required [19]. The data transformation process was then performed by removing irrelevant attributes and converting categorical variables into numerical formats to meet the requirements of machine learning algorithms [20].

2.3. K-Fold Cross Validation

K-Fold Cross Validation is a model evaluation technique that divides a dataset into k equally sized folds so that each data instance is used for both training and testing [21]. This approach reduces the bias that may arise from a single train-test split, producing a more reliable and stable performance assessment. The value of k is determined from the relationship between the total data size (N) and the fold size (α).

2.4. Student Performance and Exam Scores

Student performance is a critical indicator of academic quality and is often defined by scores or grades on various assessments, culminating in the final exam score [22]. The prediction and classification of student

performance are essential objectives in Educational Data Mining (EDM), as successful models enable institutions to identify at-risk students and implement targeted interventions [22].

The classification model proposed in this study focuses on predicting student performance based on several measurable input factors. Prior research consistently highlights several variables as significant predictors of a student's final outcome:

Factors Influencing Exam Scores

1. **Hours Studied:** The amount of time dedicated to studying is a direct proxy for student effort and is empirically correlated with positive learning outcomes. Studies utilizing decision trees and other machine learning algorithms often rank study time as an important feature for predicting grades [23].
2. **Attendance:** Student attendance, often represented as a percentage of classes or lectures attended, reflects engagement and consistent exposure to course material. High attendance rates are generally associated with better performance and are frequently cited as a critical predictor in various classification models of student performance [23].
3. **Previous Scores:** A student's academic history, represented by previous scores, is a highly reliable indicator of future performance. Past success suggests a baseline level of knowledge retention, study habits, and overall aptitude, making it a crucial feature in both ensemble models, such as Random Forest, and kernel-based models, such as SVM, for predicting final grades [24].
4. **Sleep Hours:** While often overlooked, recent literature has explored the non-academic factors influencing performance. Sleep hours are integral to cognitive function, memory consolidation, and overall well-being. Adequate sleep is necessary for optimal academic functioning and can be a contributing factor in models designed to classify academic success [24].

The classification approach used in this research employs machine learning algorithms (Decision Tree, Random Forest, and SVM) to model the relationship between these input factors and the target variable, Exam Score. By classifying the Exam Scores into distinct categories (e.g., 'Pass' or 'Fail'), the study aims to provide a robust, interpretable model for assessing student academic status.

2.5. Classification

The Classification Process is a crucial stage in which the performance of various predictive algorithms is evaluated and compared [25]. The study focuses on evaluating two major machine learning algorithms: Decision Tree and Random Forest, which are considered appropriate for classifying student exam scores.

2.6. Confusion Matrix

The Confusion Matrix is one of the most widely used tools for evaluating classification performance, as it compares predicted and actual class labels to summarize outcomes in terms of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). From these four components, the key performance indicators used in this study, accuracy, precision, and recall, are derived. Table 1 presents the structure of the confusion matrix adopted in this research.

Table 1. Confusion Matrix

		Prediction	
		Positive	Negative
Actual	Positive	TP	FN
	Negative	FP	TN

The computation of accuracy follows equation (1); while precision and recall are calculated using equations (2) and (3), in that order [26].

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

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

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

2.7. Decision Tree

The Decision Tree algorithm is widely adopted in Educational Data Mining (EDM) for its strong predictive power and interpretability [27]. Decision Trees classify data by recursively partitioning the dataset

into purer subsets. At each node, the algorithm selects the feature and split point that minimizes impurity (or maximizes information gain).

One of the most common measures of impurity used for classification is Entropy, which quantifies the disorder or randomness of class labels within a node and is calculated as Equation 4.

$$\text{Entropy}(D) = - \sum_{k=1}^m P_k \log_2 P_k \quad (5)$$

The goal of the Decision Tree algorithm is to find the split that produces the greatest reduction in Entropy (i.e., the highest Information Gain) across the resulting child nodes, leading to more homogeneous groups.

2.8. Random Forest

Random Forest is an ensemble learning method that constructs multiple Decision Trees, each trained on a bootstrap sample of the data with a randomly selected subset of features at each split, and aggregates their predictions via majority voting to produce a more accurate and stable classification [28]. This bagging mechanism reduces the variance associated with individual trees and makes Random Forest more resistant to overfitting than a single Decision Tree. Prior studies confirm that Random Forest consistently achieves higher accuracy than a single Decision Tree across various educational classification tasks, including student grade prediction.

$$H(x) = \arg \max \sum_{k=1}^k I(h_k(x) = c) \quad (6)$$

2.9. Literature Review

Student performance prediction is a central topic in Educational Data Mining (EDM) and is commonly formulated as a classification problem when learning outcomes are represented in discrete categories, such as grades or pass/fail status. Sivakumar and Venkataraman demonstrated that supervised learning models, particularly ensemble-based approaches such as Random Forest, are effective for grade classification tasks and offer strong predictive performance in educational datasets [29]. Furthermore, Zhang et al. provided a comprehensive review of EDM techniques and highlighted that classification-based models are widely used to predict academic performance using demographic, academic, and behavioral features. Their study emphasized that supervised learning models, including decision-tree-based approaches, are effective for identifying students at risk and supporting academic decision-making [30].

Decision Tree algorithms are frequently applied in educational data analysis due to their simplicity and interpretability. Dahal and Shakya compared several decision tree algorithms, including ID3, C4.5, and CART, with Random Forest and other classifiers for predicting student results. Their findings indicate that decision tree models perform well on small to medium-sized datasets and are particularly effective for binary and multi-class classification of student exam outcomes. The study also highlights that decision trees are easily interpretable, making them suitable for use by educators and academic administrators [31].

Similarly, Huynh-Cam et al. employed decision tree algorithms to predict first-year university students' learning performance using pre-semester data. The results showed that CART achieved competitive accuracy while also identifying key contributing factors such as parental occupation and financial background. This demonstrates the strength of decision trees not only in prediction accuracy but also in extracting meaningful decision rules from educational data [32]. Several studies have highlighted the trade-off between interpretability and predictive performance when comparing Decision Tree and Random Forest models. Dahal and Shakya found that while decision trees offer clearer interpretability, Random Forest generally provides higher accuracy due to its ensemble structure [31]. Huynh-Cam et al. further confirmed that Random Forest tends to produce more stable predictions, whereas decision trees excel in explaining the underlying factors influencing student performance [32].

Based on these findings, a direct comparative analysis between the Decision Tree and Random Forest models remains relevant, particularly for student exam score classification tasks that require both accurate predictions and explainability. Although previous studies have examined Decision Tree and Random Forest models in educational contexts, most research focuses on comparing Random Forest with other algorithms or evaluating multiple classifiers simultaneously [29], [33]. There is still limited research that specifically conducts a focused comparative analysis between Decision Tree and Random Forest models for student exam score classification using a consistent dataset and evaluation framework. Therefore, this study aims to address this gap by systematically comparing Decision Tree and Random Forest models to determine their effectiveness in classifying student exam scores.

3. RESULT AND DISCUSSION

3.1. Initial Data

The dataset used in this study consists of six attributes describing student characteristics and academically related behaviors. Student_id is a categorical attribute used solely as a unique identifier and does not influence the predictive model. Hours_studied represents the average amount of time a student spends studying, while Sleep_hours indicates the number of hours the student sleeps, both of which are numerical factors that may impact learning outcomes. Attendance_percent measures a student's class attendance and reflects their engagement in the learning process. Previous_scores captures the student's academic history through past assessment results and serves as a strong predictor of future performance. Lastly, Exam_scores is the numerical target variable representing the student's final exam result, which the classification model aims to predict based on the other attributes. Table 2 shows the details of each attribute.

Table 2. Initial Data

No	Attribute Name	Data Type	Description
1	Student_id	Categorical	A unique identifier for each student.
2	Hours_studied	Numerical	The average number of hours the student spent studying per period (e.g., per day/week)
3	Sleep_hours	Numerical	Number of hours the student sleeps
4	Attendance_percent	Numerical	The student's class or relevant session attendance percentage.
5	Previous_scores	Numerical	Score or mark from prior exams/assessments.
6	Exam_scores	Numerical	Score of the student's exam

3.2. Data Preprocessing

As part of the data preprocessing stage, the raw dataset was cleaned, normalized, and transformed to ensure consistency and improve model performance. The table below presents a sample of the preprocessed data after unnecessary values were removed and numerical features were standardized. This refined dataset includes key attributes such as hours studied, attendance percentage, previous scores, and exam scores, which have been processed to ensure each variable is in a suitable format for machine-learning analysis. The structured output shown in the table reflects the final form of the data before it is used for model training and evaluation. Table 3 presents the dataset in the preprocessing step.

Table 3. Preprocessing Data Result

No	Student_id	Hours_studied	Attendance_percent	Previous_scores	Exam_scores
1	6.3	6.6	74.83	66.80	33.95
2	3.2	1.4	14.24	15.66	6.7
3	1.0	4.0	50.30	40.00	17.10
4	3.5	5.3	62.20	54.00	29.50
5	6.1	6.7	75.25	67.50	34.05
6	9.0	8.0	87.42	80.00	38.75
7	12.0	9.0	100	95.00	51.30

3.3. Implementation of the Classification Algorithm

This study applies the Decision Tree and Random Forest classification algorithms in Python to classify student exam scores. To evaluate the performance and generalization capability of both models, the K-Fold Cross-Validation method with five folds was employed. The accuracy obtained in each fold and the corresponding mean accuracy are presented in Table 4.

Table 4. K-Fold Cross-Validation Accuracy Results (k=5)

Algorithm	Fold 1 (%)	Fold 2 (%)	Fold 3 (%)	Fold 4 (%)	Fold 5 (%)	Mean Accuracy (%)
Decision Tree	85.00	80.00	77.50	87.50	77.50	81.50
Random Forest	82.50	92.50	82.50	95.00	80.00	86.50

Table 4 shows that the Random Forest algorithm achieved a mean accuracy of 86.50%, while the Decision Tree algorithm obtained 81.50%. The highest accuracy for the Decision Tree model was recorded in Fold 4 (87.50%), whereas the Random Forest model achieved its highest accuracy in Fold 2 (92.50%) and Fold 4 (95.00%).

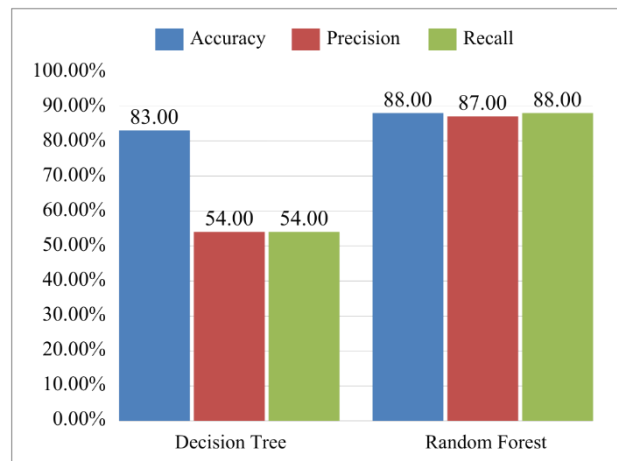


Figure 2. Comparison of Algorithm Classification Results

Figure 2 presents the classification performance of the Decision Tree and Random Forest algorithms based on accuracy, precision, and recall. The Decision Tree model achieved an accuracy of 83%, a precision of 54%, and a recall of 54%. Meanwhile, the Random Forest model achieved an accuracy of 88%, a precision of 87%, and a recall of 88%.

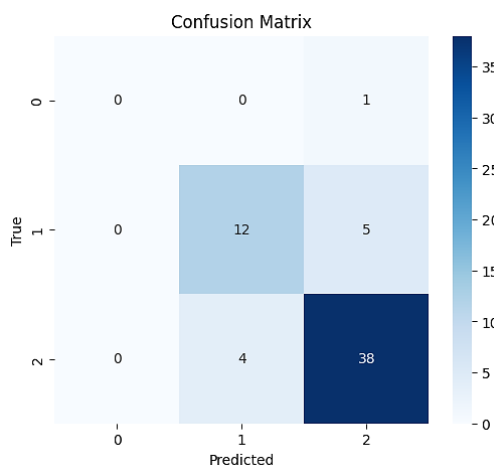


Figure 3. Confusion matrix of the Decision Tree

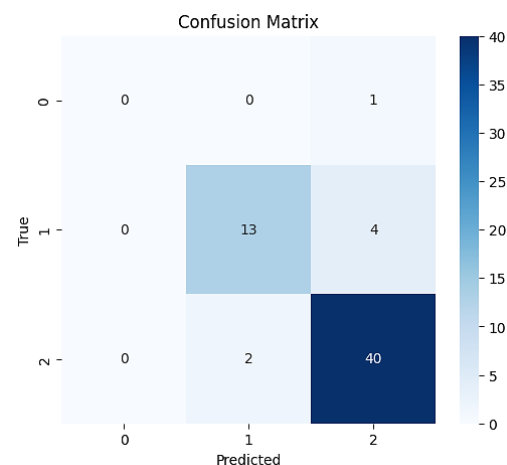


Figure 4. Confusion matrix of Random Forest

Figure 3 presents the confusion matrix of the Decision Tree algorithm. The model correctly classified 38 instances of class 2 and 12 instances of class 1. The single instance belonging to class 0 was misclassified, and several observations from class 1 were predicted as class 2. Figure 4 presents the confusion matrix of the Random Forest algorithm. The model correctly classified 40 instances of class 2 and 13 instances of class 1. Similar to the Decision Tree model, the single instance of class 0 was misclassified.

To further examine the dataset's characteristics, descriptive and feature correlation analyses were conducted on the numerical attributes (Hours_studied, Attendance_percent, Previous_scores, and Exam_scores). These analyses provide additional information regarding the distribution of the data and the relationships among variables. The interpretation of these statistical findings and their influence on the classification performance are discussed in Section 3.4.

3.4. Discussion

The experimental results demonstrate that the Random Forest algorithm consistently outperformed the Decision Tree algorithm in classifying student exam scores. Based on the 5-Fold Cross-Validation results, Random Forest achieved a mean accuracy of 86.50%, compared to 81.50% for Decision Tree. Similarly, the evaluation metrics in Figure 2 show that Random Forest achieved higher accuracy (88%), precision (87%), and recall (88%), indicating more balanced and reliable classification performance. These results suggest that the ensemble learning strategy employed by Random Forest enables the model to generalize better than a single Decision Tree.

The superior performance of Random Forest can be explained by its ability to construct multiple decision trees using different subsets of training data and randomly selected features. This ensemble mechanism reduces the risk of overfitting, which is a common limitation of Decision Tree models. In contrast, Decision Tree relies on a single tree structure, making it more sensitive to variations and noise in the training dataset. Consequently, Random Forest yields more stable predictions across validation folds and is more robust when classifying unseen data.

The characteristics of the dataset also contribute to the observed performance differences. The dataset consists primarily of numerical variables, including Hours_studied, Attendance_percent, Previous_scores, and Exam_scores, which exhibit meaningful relationships with student academic performance. Among these variables, Previous_scores shows the strongest positive correlation with Exam_scores, indicating that historical academic achievement is an important predictor of future performance. Meanwhile, Hours_studied and Attendance_percent provide complementary information regarding students' learning behavior. Because these variables interact with one another, Random Forest is better suited to capture their complex relationships than a single Decision Tree.

The confusion matrices further support these findings. Both algorithms correctly classified the majority of instances in class 2, while class 0 remained difficult to identify due to the very limited number of observations in that category. This imbalance likely reduced the models' ability to learn representative patterns for the minority class. Nevertheless, Random Forest produced fewer misclassifications between class 1 and class 2, demonstrating its improved capability to distinguish similar categories. These results indicate that the ensemble approach provides better class discrimination even when the dataset contains uneven class distributions.

The findings of this study are consistent with previous Educational Data Mining (EDM) research, which has frequently reported that Random Forest achieves higher predictive performance than a single Decision Tree because of its ensemble learning mechanism and stronger generalization capability. Similar conclusions have been reported in studies on student performance prediction, where Random Forest effectively captures nonlinear relationships among academic and behavioral variables while reducing prediction variance.

From a practical perspective, the proposed classification model can help educational institutions identify students with varying levels of academic performance at an early stage. Such predictions enable lecturers, academic advisors, and school administrators to design targeted interventions, including additional tutoring, personalized learning assistance, and academic counseling for students at risk of lower performance. Therefore, the application of Random Forest in Educational Data Mining provides valuable support for evidence-based academic decision-making and helps improve student learning outcomes.

4. CONCLUSION

This research aimed to analyze the comparative performance of the Decision Tree and Random Forest algorithms in classifying student exam scores, contributing to data-driven decision-making in education. The study employed a systematic approach comprising data preprocessing, K-Fold Cross-Validation, and evaluation with performance metrics derived from the Confusion Matrix.

The comparative analysis decisively concludes that the Random Forest algorithm offers superior overall predictive performance and enhanced stability for student exam score classification. Specifically, the Random Forest model achieved a Mean Accuracy of 86.50% across the 5-Fold Cross-Validation, significantly surpassing the Decision Tree algorithm, which recorded an average accuracy of 81.50%. Furthermore, the analysis of the Confusion Matrices confirms Random Forest's superior generalization ability, as evidenced by its successful classification of 40 instances in Class 2 and 13 instances in Class 1, resulting in fewer misclassifications compared to the Decision Tree. Consequently, the derived Random Forest model provides a robust and accurate predictive framework that educational administrators can use to identify students' performance status, enabling targeted intervention strategies to support at-risk students. Future work should extend the comparative analysis by integrating additional high-performing machine learning algorithms frequently cited in the Educational Data Mining (EDM) literature, such as Support Vector Machines (SVMs) and Gradient Boosted Trees (XGBoost), to establish a broader benchmark for exam score classification. Furthermore, to address the poor classification performance observed for the minority category (Class 0), advanced data-handling techniques, including oversampling (e.g., SMOTE) or undersampling, must be applied during data preprocessing to mitigate class imbalance. Finally, an in-depth feature-importance analysis should be conducted on the superior Random Forest model to precisely quantify and rank the contributions of key input factors, such as Hours Studied and Previous Scores, thereby providing actionable insights into the key determinants of student academic success.

REFERENCES

- [1] Ahmed, E. (2024). Student Performance Prediction Using Machine Learning Algorithms. *Applied Computational Intelligence and Soft Computing*, 2024, Article 4067721. <https://doi.org/10.1155/2024/4067721>

- [2] Akhatkulov, S., Yusupov, O., & Omonov, A. (2024). Predicting students' future final exam results using machine learning algorithms: A comparative analysis. *AIP Conference Proceedings*, 3244(1), 030071.
- [3] Al-Din, M. S. N., & Al Abdulqader, H. A. (2024). Students' Academic Performance Prediction Using Educational Data Mining and Machine Learning: A Systematic Review. *International Journal of Research and Innovation in Social Science*, 8(8), 1264–1291. <https://doi.org/10.47772/IJRISS.2024.808095>
- [4] Al-Tameemi, G., Xue, J., Hadi, I., & Ajit, S. (2024). A Hybrid Machine Learning Approach for Predicting Student Performance Using Multi-class Educational Datasets. *Procedia Computer Science*, 238, 888–895. <https://doi.org/10.1016/j.procs.2024.06.108>
- [5] Assegie, T. A., Salau, A. O., Chhabra, G., Kaushik, K., & Braide, S. L. (2024). Evaluation of Random Forest and Support Vector Machine Models in Educational Data Mining. *IEEE*.
- [6] Badal, Y. T., & Sungkur, R. K. (2023). Predictive Modelling and Analytics of Students' Grades Using Machine Learning Algorithms. *Education and Information Technologies*, 28, 3027–3057. <https://doi.org/10.1007/s10639-022-11299-8>
- [7] Chen, Y., & Jin, K. (2024). Educational Performance Prediction with Random Forest and Innovative Optimizers: A Data Mining Approach. *International Journal of Advanced Computer Science and Applications*, 15(3). <https://doi.org/10.14569/IJACSA.2024.0150308>
- [8] Helmud, E., Fitriyani, F., & Romadiana, P. (2024). Classification Comparison Performance of Supervised Machine Learning Random Forest and Decision Tree Algorithms Using Confusion Matrix. *Jurnal SISFOKOM*.
- [9] Hussain, M. M., Akbar, S., Hassan, S. A., Aziz, M. W., & Urooj, F. (2024). Prediction of Student's Academic Performance through Data Mining Approach. *Journal of Informatics and Web Engineering*, 3(1). <https://doi.org/10.33093/jiwe.2024.3.1.16>
- [10] Ibrahimi, E., Lopes, M. B., Dharmo, X., et al. (2023). Overview of Data Preprocessing for Machine Learning Applications in Human Microbiome Research. *Frontiers in Microbiology*, 14, 1250909. <https://doi.org/10.3389/fmicb.2023.1250909>
- [11] Kurniasari, D., Hidayah, R. N., Notiragayu, N., Warsono, & Nisa, R. K. (2024). Classification Models for Academic Performance: A Comparative Study of Naïve Bayes and Random Forest Algorithms in Analyzing University Student Grades. *Jurnal Teknik Informatika (JUTIF)*, 5(5), 1267–1276. <https://doi.org/10.52436/1.jutif.2024.5.5.2066>
- [12] Kumar, M., Singh, N., Wadhwa, J., Singh, P., Kumar, G., & Qtaishat, A. (2024). Utilizing Random Forest and XGBoost Data Mining Algorithms for Anticipating Students' Academic Performance. *International Journal of Modern Education and Computer Science*, 16(2), 29–44. <https://doi.org/10.5815/ijmecs.2024.02.03>
- [13] Nachouki, M., Mohamed, E. A., Mehdi, R., & Abou Naaj, M. (2023). Student Course Grade Prediction Using the Random Forest Algorithm: Analysis of Predictors' Importance. *Trends in Neuroscience and Education*.
- [14] Salloum, S. A., Salloum, A., Shaalan, K., Alfaisal, R., & Basiouni, A. (2024). Comparative Analysis of Classical Machine Learning Techniques for Predicting Students' Exam Performance. In *Communications in Computer and Information Science* (Vol. 2162, pp. 219–227). Springer. https://doi.org/10.1007/978-3-031-65996-6_19
- [15] Saeed, M. M. (2024). Forecasting the Academic Performance by Leveraging Educational Data Mining. *Intelligent Automation & Soft Computing*, 39(2), 213–231. <https://doi.org/10.32604/iasc.2024.043020>
- [16] Sivakumar, S., & Venkataraman, S. (2025). Evaluating Machine Learning Approaches: A Comparative Study of Random Forest and Neural Networks in Grade Classification. *Indonesian Journal of Data and Science*, 6(1), 73–80. <https://doi.org/10.56705/ijodas.v6i1.240>
- [17] Sombat, W., Kaensar, C., Hiranpongson, S., & Baokham, S. (2024). Student Academic Success Prediction System Using Random Forest Technique. *Journal of Science and Science Education*, 7(2), 245–259. <https://doi.org/10.14456/jsse.2024.19>
- [18] Tanveer, H., Adam, M. A., Khan, M. A., Ali, M. A., & Shakoor, A. (2023). Analyzing the Performance and Efficiency of Machine Learning Algorithms on Various Datasets and Applications. *Asian Bulletin of Big Data Management*, 3(2), 126–136. <https://doi.org/10.62019/abdbm.v3i2.83>
- [19] Ünalın, S., Günay, O., Akkurt, I., Gunoglu, K., & Tekin, H. O. (2024). A Comparative Study on Breast Cancer Classification with Stratified Shuffle Split and K-Fold Cross Validation via Ensembled Machine Learning. *Journal of Radiation Research and Applied Sciences*, 17(4), 101080. <https://doi.org/10.1016/j.jrras.2024.101080>
- [20] Zhang, Y., Yun, Y., An, R., Cui, J., Dai, H., & Shang, X. (2021). Educational Data Mining Techniques for Student Performance Prediction: Method Review and Comparison Analysis. *Frontiers in Psychology*, 12, 698490. <https://doi.org/10.3389/fpsyg.2021.698490>

- [21] Ahmed, E. (2024). Student Performance Prediction Using Machine Learning Algorithms. *Applied Computational Intelligence and Soft Computing*, 2024, Article 4067721. <https://doi.org/10.1155/2024/4067721>
- [22] Akhatkulov, S., Yusupov, O., & Omonov, A. (2024). Predicting students' future final exam results using machine learning algorithms: A comparative analysis. *AIP Conference Proceedings*, 3244(1), 030071.
- [23] Al-Din, M. S. N., & Al Abdulqader, H. A. (2024). Students' Academic Performance Prediction Using Educational Data Mining and Machine Learning: A Systematic Review. *International Journal of Research and Innovation in Social Science*, 8(8), 1264–1291. <https://doi.org/10.47772/IJRISS.2024.808095>
- [24] Chen, Y., & Jin, K. (2024). Educational Performance Prediction with Random Forest and Innovative Optimizers: A Data Mining Approach. *International Journal of Advanced Computer Science and Applications*, 15(3). <https://doi.org/10.14569/IJACSA.2024.0150308>
- [25] Helmud, E., Fitriyani, F., & Romadiana, P. (2024). Classification Comparison Performance of Supervised Machine Learning Random Forest and Decision Tree Algorithms Using Confusion Matrix. *Jurnal SISFOKOM*.
- [26] Hussain, M. M., Akbar, S., Hassan, S. A., Aziz, M. W., & Urooj, F. (2024). Prediction of Student's Academic Performance through Data Mining Approach. *Journal of Informatics and Web Engineering*, 3(1). <https://doi.org/10.33093/jiwe.2024.3.1.16>
- [27] Ibrahim, E., Lopes, M. B., Dharmo, X., Simeon, A., Shigdel, R., Hron, K., Stres, B., D'Elia, D., Berland, M., & Marcos-Zambrano, L. J. (2023). Overview of Data Preprocessing for Machine Learning Applications in Human Microbiome Research. *Frontiers in Microbiology*, 14, Article 1250909. <https://doi.org/10.3389/fmicb.2023.1250909>
- [28] Khairy, D., Alharbi, N., Amasha, M. A., Areed, M. F., Alkhalaf, S., & Abougalala, R. A. (2024). Prediction of Student Exam Performance Using Data Mining Classification Algorithms. *Education and Information Technologies*, 29, 21621–21645. <https://doi.org/10.1007/s10639-024-12619-w>
- [29] Kurniasari, D., Hidayah, R. N., Notiragayu, N., Warsono, & Nisa, R. K. (2024). Classification Models for Academic Performance: A Comparative Study of Naïve Bayes and Random Forest Algorithms in Analyzing University Student Grades. *Jurnal Teknik Informatika (JUTIF)*, 5(5), 1267–1276. <https://doi.org/10.52436/1.jutif.2024.5.5.2066>
- [30] Kumar, M., Singh, N., Wadhwa, J., Singh, P., Kumar, G., & Qtaishat, A. (2024). Utilizing Random Forest and XGBoost Data Mining Algorithms for Anticipating Students' Academic Performance. *International Journal of Modern Education and Computer Science*, 16(2), 29–44. <https://doi.org/10.5815/ijmecs.2024.02.03>
- [31] Nachouki, M., Mohamed, E. A., Mehdi, R., & Abou Naaj, M. (2023). Student Course Grade Prediction Using the Random Forest Algorithm: Analysis of Predictors' Importance. *Trends in Neuroscience and Education*.
- [32] Salloum, S. A., Salloum, A., Shaalan, K., Alfaisal, R., & Basiouni, A. (2024). Comparative Analysis of Classical Machine Learning Techniques for Predicting Students' Exam Performance. In A. Basiouni & C. Frasson (Eds.), *Communications in Computer and Information Science* (Vol. 2162, pp. 219–227). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-65996-6_19
- [33] Sivakumar, S., & Venkataraman, S. (2025). Evaluating Machine Learning Approaches: A Comparative Study of Random Forest and Neural Networks in Grade Classification. *Indonesian Journal of Data and Science*, 6(1), 73–80. <https://doi.org/10.56705/ijodas.v6i1.240>
- [34] Tanveer, H., Adam, M. A., Khan, M. A., Ali, M. A., & Shakoor, A. (2023). Analyzing the Performance and Efficiency of Machine Learning Algorithms on Various Datasets and Applications. *Asian Bulletin of Big Data Management*, 3(2), 126–136. <https://doi.org/10.62019/abbdm.v3i2.83>
- [35] Ünal, S., Günay, O., Akkurt, I., Gunoglu, K., & Tekin, H. O. (2024). A Comparative Study on Breast Cancer Classification with Stratified Shuffle Split and K-Fold Cross Validation via Ensembled Machine Learning. *Journal of Radiation Research and Applied Sciences*, 17(4), 101080. <https://doi.org/10.1016/j.jrras.2024.101080>