Semarak International Journal of Machine Learning

Journal homepage: https://semarakilmu.my/index.php/sijml/index
ISSN: 3030-5241

A Comparative Analysis of Classification Algorithms on Student Academic Performance

Tunde Mojeed Adeyanju[1,*], Azwa Abdul Aziz[2], Suhailan Safei[1]

[1] Faculty of Informatics and Computing, Universiti Sultan Zainal Abdin Besut Campus 22200, Besut Terengganu Malaysia
[2] Data Science & Analytics (DASA) Special Interest Group (SIG) Universiti Sultan Zainal Abidin (UniSZA), Gong Badak Terengganu 20300, Malaysia

ARTICLE INFO

Article history:
Received 10 February 2025
Received in revised form 10 March 2025
Accepted 20 March 2025
Available online 21 March 2025

ABSTRACT

Understanding the best use of machine learning algorithms, such as decision tree classifiers, in predicting academic performance is crucial for educational institutions to identify at-risk students and implement targeted interventions. It is essential to determine the best algorithm for academic achievement to ensure accurate predictions and effective support mechanisms for students. This study investigated the multifaceted factors influencing academic performance in computer science research, emphasizing the significance of benchmarking prediction and classification algorithms such as decision trees, random forest, AdaBoost, K nearest neighbors, and support vector machines for discerning optimal models. After classifying the ML algorithms and their implications for student academic performance, we rigorously evaluated algorithm performance through k-fold cross-validation and training-testing splits. The benchmarking process is critical for assessing the efficacy of machine learning algorithms in accurately predicting academic performance, providing a foundation for informed decision-making in educational interventions and policy development. This study underscores the importance of benchmarking machine learning algorithms as a fundamental step in assessing their suitability for predicting academic performance, thereby laying the groundwork for effective educational strategies and interventions. While random forest emerged as the top performer with a mean accuracy of 75.88%, decision trees, as the primary algorithm for our research, underwent further optimization using GridSearchCV. This results in a notable mean accuracy of 76.31%, highlighting the efficacy of the decision tree classifier in understanding the complex dynamics impacting academic performance. These findings offer valuable insights for enhancing student achievement across diverse demographic conditions. This research illustrates that explainable machine learning models retain their importance due to their transparency and computational efficiency. The outcomes offer practical recommendations for policymakers, advocating for early intervention programs, personalized learning strategies, and effective resource distribution. Future investigations should consider hybrid models that incorporate deep learning while ensuring interpretability to further improve predictions of student success.

Keywords:
Student academic performance; machine learning; decision tree; socioeconomic status

* *Corresponding author.*
E-mail address: teeboy470@yahoo.com

https://doi.org/10.37934/sijml.5.1.7486a

1. Introduction

Studying the factors influencing student performance is essential for educational institutions to implement targeted interventions and support mechanisms that improve student outcomes. By identifying and addressing disparities and understanding the factors that influence academic performance, educators and policymakers can implement evidence-based practices and policies that promote equity, inclusion, and student success across diverse student populations. Investigating the multifaceted factors affecting academic performance is critical for enhancing educational outcomes and ensuring the academic success of all students.

This research aimed to address these limitations and improve the accuracy of predicting academic performance among students. Specifically, the study evaluates the role of demographic and socioeconomic factors, including parental education, family income, and study habits, in shaping student achievement. By leveraging machine learning models, this research seeks to identify the most effective classification algorithms for academic performance prediction while ensuring that the models remain interpretable for practical implementation in educational settings. This objective is driven by the need to overcome challenges related to feature selection and algorithm choices, as highlighted in previous studies [5,9].

While deep learning models have gained popularity, their interpretability remains a challenge in educational applications, where transparency is crucial. This study focuses on classical machine learning models, particularly Decision Trees and ensemble methods, to provide a balance between performance and transparency. By benchmarking these models using data sourced from the UCI Machine Learning Repository, this research provides a robust foundation for data-driven decision-making in education. Machine learning algorithms provide an opportunity to develop predictive models that help educators identify at-risk students and implement personalized interventions.

Existing studies have primarily focused on hybrid approaches combining Decision Trees with neural networks or support vector machines [8,7]. However, there is a lack of benchmarking studies that comprehensively compare the effectiveness of classical machine learning models in academic performance prediction. This research aims to fill that gap by benchmarking key machine-learning classifiers on a dataset sourced from the UCI Machine Learning Repository.

2. Literature Review

This research aimed to address these limitations and improve the accuracy of predicting academic performance among students. Specifically, the research evaluates factors and algorithms that may enhance accuracy, particularly in the context of student demographics and family dynamics. By addressing challenges related to feature selection, SES factors, and algorithm choices, as highlighted in the studies by Shyama *et al.,* [5] and Chen *et al.,* [9], this study provides a robust benchmarking analysis.

To better position our study, Table 1 provides a comparison of existing research on academic performance prediction. Various studies have explored the use of machine learning algorithms to predict student academic performance. Shyama *et al.,* [5] analysed student demographics and academic achievement, emphasizing the role of socioeconomic status and family dynamics. Chen *et al.,* [9] reviewed decision tree applications in educational data mining, noting their effectiveness in predicting student outcomes. Wen *et al.,* [8] explored the application of deep neural networks for forecasting student performance by analyzing sequences of online learning activities. Their research revealed that deep learning methodologies can accurately model temporal learning patterns, thereby enhancing early intervention approaches. Nonetheless, a significant limitation identified was the necessity for additional validation across a variety of datasets to improve the generalizability of the findings. H. Alamri *et al.,* [7] explored the application of hybrid models that integrate Decision Trees with Support Vector Machines (SVM) to forecast student academic outcomes. Their findings indicated that this integration improves classification accuracy by utilizing the interpretative capabilities of decision trees alongside the robust classification power of SVM. But faced limitations due to its reliance on a single dataset, which constrained its generalizability, and lack of thorough feature selection methodology. Significant challenges in their hybrid model that combines decision trees with fuzzy genetic algorithms [12]. Although their methodology enhanced prediction accuracy and the robustness of the model, they observed constraints related to scalability and computational efficiency. The optimisation process inherent in the fuzzy genetic algorithm contributed to increased computational complexity, rendering it less suitable for extensive datasets. Furthermore, the model's ability to adapt to various educational settings was identified as a concern, necessitating additional adjustments for different academic environments. These limitations indicate a need to develop more efficient hybrid techniques or feature selection strategies to improve classification performance while ensuring scalability in predicting student academic outcomes.

These studies underscore the need for benchmarking traditional machine learning models to balance predictive accuracy, computational efficiency, and interpretability, particularly in educational settings where transparency is crucial for policy implementation. Munir *et al.,* [10] studied punctuality and participation in other activities as key factors in predicting academic achievements. They noted that prediction accuracy could be enhanced if socioeconomic status (SES) factors were included and analyzed in more detail. The integration of SES could provide deeper insights into students' backgrounds, thereby improving the accuracy of predictive models.

Additionally, [2] examined demographic data, family background, and academic performance, revealing that SES influences academic outcomes but is not fully mediated by commonly used variables. They suggested that better calibration of SES indicators, such as parental education and occupation, may enhance the predictive power of models. Similarly, [3] acknowledged the importance of SES but found its contribution to academic performance prediction to be unclear, recommending further exploration of its combined effect with other factors.

Huang *et al.,* [4] investigated prior academic performance, attitudes toward learning, and home background, but their study indicated that predictive accuracy was limited due to insufficient SES data. Past scores, study habits, and family background, concluding that the absence of extensive SES data undermined the predictive potential of models [6]. They recommended that future studies examine the relationship between SES and other influential factors to refine prediction models [6].

All these researchers have noted the importance of SES inclusion in predictive models to provide a more holistic view of students' circumstances and potentially lead to more accurate predictions. Unfortunately, studies on SES among the factors that predict academic achievement are lacking.

Table 1
Comparison of existing research on academic performance prediction

Study	Algorithms Used	Dataset Source	Key Findings	Gap Identified
Chen *et al.,* [9]	Decision Tree	Public Dataset	Application of decision tree algorithm in educational data mining	Need for hybrid approaches and deeper SES analysis
Shyama *et al.,* [5]	Random Forest, SVM	Public Dataset	SES factors improve predictive performance	More refined SES indicators needed
Wen *et al.,* [8]	Deep Neural Network	Online Learning Activity Sequence	Early prediction of student performance	Need for further validation on different datasets
Aslam *et al.,* [7]	Decision Tree + SVM	Public Dataset	Predicting student academic performance using hybrid models	Further validation needed on larger datasets
Hamsa *et al.,* [12]	Decision Tree + Fuzzy Genetic Algorithm	Public Dataset	Improved prediction accuracy and model robustness	Scalability and computational complexity issues
Zhang *et al.,* [10]	Decision Tree	Public Dataset	Participation and punctuality improve predictions	SES should be integrated for better accuracy
Rodríguez-Hernández *et al.,* [2]	Various ML Models	Public Dataset	SES has an impact but needs better calibration	Need for refined SES measures to enhance model performance
Lutfiu *et al.,* [3]	Correlational Analysis	Public Dataset	SES positively impacts academic performance	Lacks advanced predictive modeling for comprehensive analysis
Huang *et al.,* [4]	Regression Models	Public Dataset	Prior academic performance and home background are predictive	Predictive ability limited due to inadequate SES data
Rabia *et al.,* [6]	Decision Tree,Logistic Regression	Private Dataset	Study habits and past scores are strong indicators	Lack of SES data weakens prediction accuracy
This Study	Decision Tree,Random Forest,AdaBoost,KNN,SVM	UCI Student Performance Dataset	Benchmarking classical models, focusing on SES factors and model interpretability	Lack of prior studies incorporating SES in ML-based academic predictions

3. Methodology

The proposed method, as shown in Figure 1, aims to investigate the factors influencing academic performance among students, focusing on demographics and family dynamics. The methodology involves data collection, preprocessing, exploratory data analysis (EDA), feature selection, model building, and model evaluation. Spotcheck algorithms such as linear regression, logistic regression, support vector machines, k-nearest neighbors, and random forest algorithms will be compared with decision tree algorithm for predicting academic performance. The expected outcomes include the identification of key factors influencing academic performance, insights into model performance, and recommendations for educational institutions.

Figure 1 is a flow diagram illustrating the proposed methodology for investigating factors influencing academic performance and comparing machine learning models. The process begins with data collection, followed by data preprocessing to prepare the dataset for analysis. Exploratory data analysis (EDA) was subsequently conducted to gain insights into the data, followed by feature selection to identify relevant predictors of academic performance. Next, machine learning models are built using the selected features, and their performance is evaluated. Finally, the outcomes are analysed, conclusions are drawn, and recommendations are provided based on the findings of the study.

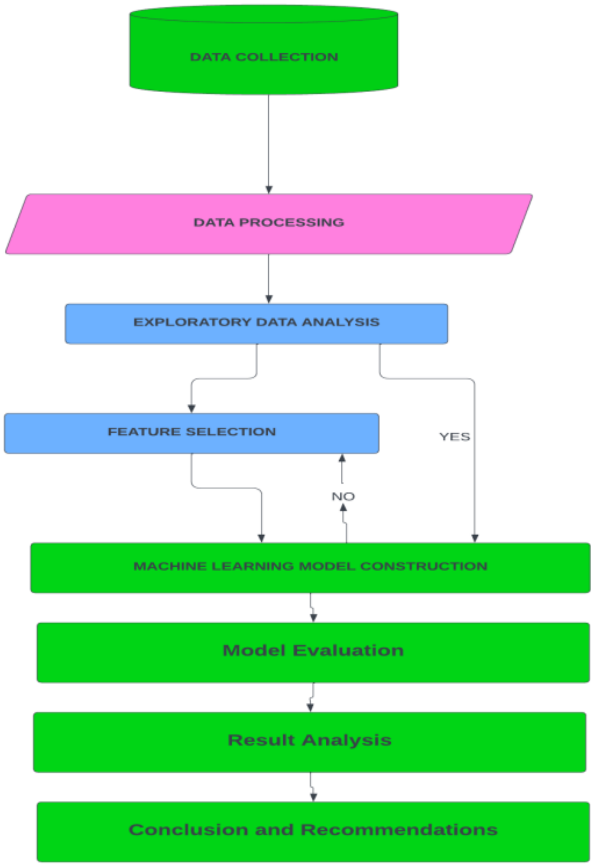

Fig. 1. Framework of benchmarking prediction Algorithm

The relevant data including demographic information, family dynamics, academic records, and other pertinent variables were gathered based on factors that may influence academic performance. These data were obtained from the UCI Machine Learning Repository, specifically the 'Student Performance ‘Dataset' [13], which includes information on students' backgrounds, social factors, and academic performance. The dataset is available at the UCI Machine Learning Repository (https://archive.ics.uci.edu/ml/datasets/Student+Performance) and was last updated in 2019. The data preparation process involved collecting data from two CSV files: 'student-mat.csv' and 'student-por.csv.'

The methodology involves data preprocessing, feature selection, model training, and performance evaluation. These collected data are then preprocessed to handle missing values, normalize or standardize the data, and encode categorical variables to prepare them for analysis. The dataset was cleaned and refined by creating a new feature, 'Final Score', which is a weighted average of three exam periods (G1, G2, G3). This step ensured that the dataset accurately reflected students' overall academic performance.

Exploratory Data Analysis (EDA) was conducted to visualize the distribution of grades across different variables, such as school, sex, age, address, family size, parental status, education, and job.

Bar plots and violin plots were generated to provide a clear and insightful visual representation of the data, aiding in the identification of patterns and anomalies on student performance scores, helping in feature selection.

To further illustrate these analyses, Figure 2. presents a bar plot showing the distribution of students' final grades across varying study durations and parental education level.

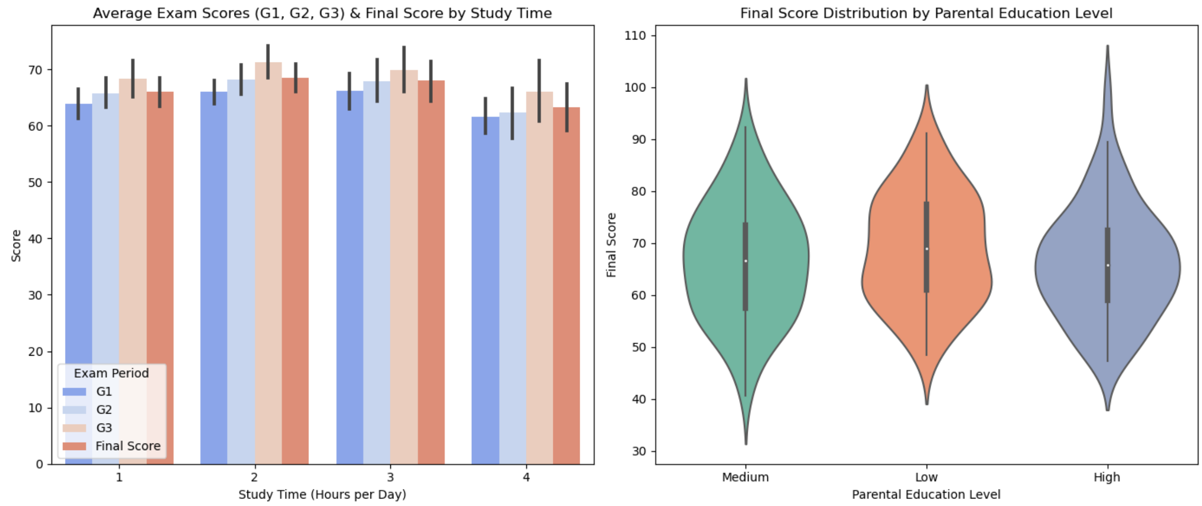

Fig. 2. Distribution of students' final grades across varying study durations

The bar chart presents the average scores for G1, G2, G3, and the calculated Final Score across varying study durations. It indicates that students who dedicate more time to their studies typically attain higher scores, with a noticeable trend where scores in later examination periods (G3) frequently surpass those in earlier periods (G1).

The violin plot depicts the distribution of Final Scores in relation to different levels of parental education. It demonstrates that students whose parents have attained higher levels of education generally achieve better Final Scores. Conversely, students from families with lower educational backgrounds display a broader range of scores, suggesting a greater variability in their academic outcomes. These visualizations highlight how different factors impact student performance and provide insights that aid in model feature selection.

Feature engineering played a crucial role in enhancing the model's predictive capabilities. New features such as 'Final Score' (a weighted average of G1, G2, and G3) and socioeconomic status (SES) indicators were introduced to capture additional context. These engineered features improved the model's ability to generalize patterns in academic performance, leading to more robust predictions.

Table 2 presents the preprocessing and feature engineering procedures that were crucial for dataset preparation prior to model training. To address missing values, mean or mode imputation techniques were employed, thereby enhancing data completeness. To ensure uniformity among numerical features, Min-Max scaling was utilized, whereas categorical variables underwent transformation via one-hot encoding. The process of feature engineering was particularly important, as it introduced the 'Final Score,' which is a weighted average of G1, G2, and G3, aimed at more accurately reflecting overall academic performance

Table 2
The preprocessing and feature engineering steps are summarized in this

Step	Description
Missing Value Handling	Replaced missing values using mean/mode imputation where necessary
Normalization/Standardization	Applied Min-Max scaling for numerical fesures to ensure uniformity
Categorical Encoding	Converted categorical variables into numerical form using one hot encoding
Feature Engineering	Created 'Final Score as a weighted average of G1,G2, and G3
Data Visualization	Bar plots and Violin plots to explore distributions and detect anomalies

These preprocessing steps ensured data quality, making the dataset suitable for machine learning analysis. The dataset consists of demographic variables, family dynamics, and academic records. Feature selection was conducted using ANOVA and recursive feature elimination (RFE) to improve model performance.

i. Demographic Variables: Such as age, sex, and address, which provide a basic profile of each student.
ii. Family Dynamics: Including parental education, parental job, and family income; these variables serve as proxies for socioeconomic status (SES) and are known to influence academic performance.
iii. Academic Records: Consisting of exam scores from three different periods (G1, G2, and G3).

To enhance predictive performance, two critical processes were applied:

i. Feature Engineering:
   - A feature, "Final Score"was created as a weighted average of G1,G2,and G3 (using a simple average) $\text{Final Score} = \frac{\text{G1+G2+G3}}{3}$. This composite score provides an overall measure of student academic achievement.
   - SES indicators were derived by aggregating variables such as parental education,,family income.these engineered features offer a holistic view of each student's socioeconomic background, which is vital for understanding performance disparities.

ii. Feature Selection:
   - ANOVA (Analysis of Variance) was used to statistically determine which features significantly affect academic performance.
   - Recursive Feature Elimination (RFE) was applied to iteratively remove less important features, ensuring that only the most relevant predictors remain. This process helps reduce noise and overfitting, ultimately improving model generalization.

The following parameters were optimized using GridSearchCV,an hyperparameter tuning technique that systematically searches through specified parameter values to find the best combination for model performance. This process helps improve the model's predictive accuracy and generalizability by preventing underfitting or overfitting. GridSearchCV was applied to optimize the following parameters for each algorithm. Table 3 displays the outcomes of hyperparameter tuning conducted with GridSearchCV to improve model performance. Essential parameters were fine-tuned for each algorithm to avoid issues of underfitting and overfitting. The optimal combinations of parameters, including maximum depth for the Decision Tree and the number of estimators for both

Random Forest and AdaBoost, resulted in enhanced accuracy. The Decision Tree recorded the highest accuracy at 76.31%, followed closely by Random Forest, while SVM exhibited the lowest accuracy.

Table 3
Hyperparameter tuning and Model Performance

Model	Hyperparameters Tuning	Best Parameters	Accuracy (%)
Decision Tree	Max Depth (d), Criterion (Gini/Entropy)	d =10, criterion= Gini	76.31
Random Forest	N_estimators (n), Max Dept (d)	n=100, d =10	75.88
AdaBoost	N_estimators (n), Learning Rate (α)	N=100,α =0.1	74.52
KNN	N_neighbors (k), Weights (Uniform/Distance)	K=5, weights =Distance	72.40
SVM	Kernel (Linear,RBF), C (Regularization)	Kernel=RBF, C=1	70.68

3.1 Hyperparameter Tuning for Decision Tree

- Decision Tree: Max depth = [5, 10, 15], Criterion = [Gini, Entropy].
- Max Depth: [5, 10, 15] → This controls the maximum depth of the tree, preventing overfitting if too deep and underfitting if too shallow.
- Criterion: [Gini, Entropy] → These define how the tree splits data. Gini measures impurity, while Entropy considers information gain.

The dataset used includes student demographic variables, family background, and academic records. The decision tree model was tuned using GridSearchCV, iterating through different max depth values and criteria. For example, when tested on the UCI Student Performance dataset, a max depth of 10 and Gini impurity criterion yielded the best accuracy (76.31%), ensuring an optimal balance between complexity and performance.

3.2 Hyperparameter Tuning for Random Forest

Random Forest builds multiple decision trees and averages their predictions to improve generalization. The following parameters were optimized:
- N_estimators: [50, 100, 200] → Number of trees in the forest. A higher number improves stability but increases computational cost.
- Max_depth: [5, 10, 15] → Limits tree depth to prevent overfitting.

On the UCI dataset, N_estimators = 100 and max_depth = 10 provided the highest accuracy of 75.88%. Random Forest outperformed AdaBoost but remained slightly less accurate than the optimized Decision Tree.

3.3 Hyperparameter Tuning for Support Vector Machine (SVM)

SVM finds an optimal hyperplane for classification. The key hyperparameters tuned were:
- Kernel: [Linear, RBF] → Defines the decision boundary shape. RBF (Radial Basis Function) helps when data is not linearly separable.
- C: [0.1, 1, 10] → Regularization parameter controlling margin size and misclassification tolerance.

The best configuration was Kernel = RBF and C = 1, yielding an accuracy of 70.68%. While SVM is effective for high-dimensional spaces, its performance lagged behind tree-based models due to the dataset's categorical nature.

3.4 Hyperparameter Tuning for AdaBoost

AdaBoost (Adaptive Boosting) is an ensemble learning method that improves weak learners by focusing more on misclassified instances. The following parameters were optimized using GridSearchCV:N_estimators: [50, 100, 200] → Determines the number of weak learners (decision stumps) to combine. More estimators can improve performance but may lead to overfitting.

Learning_rate: [0.01, 0.1, 1.0] → Controls the contribution of each weak learner. A lower learning rate requires more estimators to achieve high accuracy.

AdaBoost was trained using decision stumps as base learners on the UCI Student Performance dataset. Results showed that N_estimators = 100 and learning_rate = 0.1 provided the best trade-off between bias and variance, achieving an accuracy of 74.52%. However, while AdaBoost improved upon simpler models like KNN, it underperformed compared to ensemble methods like Random Forest due to its sensitivity to noisy data.

3.5 Hyperparameter Tuning for K-Nearest Neighbors (KNN)

KNN classifies based on the majority vote of nearest neighbors. The following were tuned:

- N_neighbors: [3,5,7] → Number of neighbors to consider. Too few cause high variance, while too many smooth out decision boundaries.
- Weights: [Uniform, Distance] → Distance-weighted neighbors improve handling of imbalanced data.
- N_neighbors = 5 and Distance weighting yielded 72.40% accuracy. KNN was computationally expensive for large datasets and sensitive to irrelevant features.

4. Result

Models Performance was measured using accuracy, precision, recall, and F1-score, as detailed in Table 4 and Figure 3.

Table 4
Effectiveness Results of the Prediction Model based on specific Algorithm

Model Name	Accuracy	Precision	Recall	F1-score
Decision Tree	76.31%	0.88	0.88	0.88
Random Forest	75.88%	0.87	0.88	0.87
AdaBoost	74.52%	0.85	0.86	0.85
KNN	72.40%	0.83	0.84	0.83
SVM	70.68%	0.81	0.82	0.81

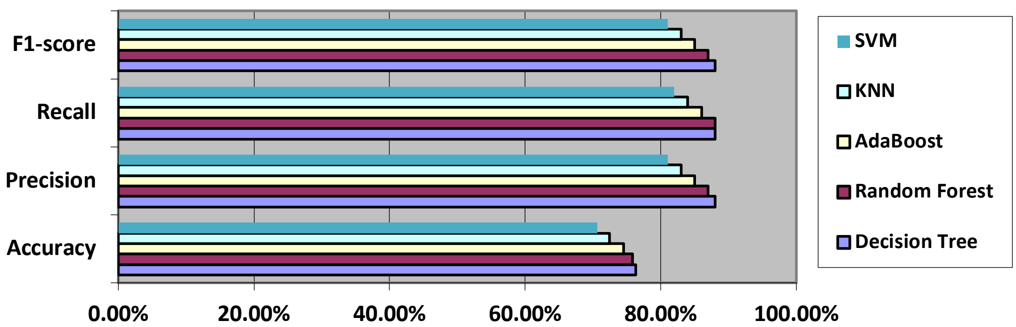

Fig. 3. Performance Comparison of Prediction Models

- Accuracy: Represents the proportion of correct predictions across all classes. The Decision Tree model achieved the highest accuracy (76.31%), indicating it had the best overall performance.
- Precision: Measures the percentage of correctly classified positive instances out of all predicted positives. The Decision Tree (0.88) and Random Forest (0.87) had the highest precision, showing they made fewer false-positive errors.
- Recall: Captures the proportion of actual positives correctly identified. Both Decision Tree and Random Forest achieved a recall of 0.88, demonstrating strong sensitivity in detecting students at different performance levels.
- F1-score: Balances precision and recall, serving as a comprehensive measure of a model's performance. The Decision Tree had the highest F1-score (0.88), confirming it as the most balanced model in this study.

These metrics provide a holistic evaluation of model performance, particularly important in cases where class imbalances may exist. The small variations between the F1-scores of different models (ranging from 0.81 to 0.88) indicate that while Decision Tree performed slightly better, the differences among top-performing models were relatively small. This highlights the need for further statistical validation, which was performed using the ANOVA (Analysis of Variance) test.

ANOVA was conducted to compare the mean accuracy scores of Decision Tree, Random Forest, AdaBoost, K-Nearest Neighbors, and Support Vector Machines to determine whether the observed performance differences were statistically significant. The test evaluates whether there are meaningful variations between multiple groups by analyzing their variance. A p-value threshold of 0.05 was used, ensuring that any observed performance differences were not due to random variation. The results indicated that Decision Tree's superior accuracy (76.31%) was statistically significant compared to the other models, validating its effectiveness in predicting student academic performance.

To further analyze the dataset, an Exploratory Data Analysis (EDA) was performed to gain insights into its structure and distribution. EDA involved visualizing key variables such as student demographics, family background, study time, and academic records to detect patterns, missing values, and potential outliers.

Descriptive statistics such as mean, median, and standard deviation were calculated to summarize numerical features, while categorical features were analyzed using frequency distributions. Correlation heatmaps were generated to identify relationships between independent variables and student performance scores.

For instance, in this study, the accuracy scores of the Decision Tree model and the Random Forest model were compared using five test folds:

- Decision Tree Accuracy: [76.5, 76.2, 75.8, 76.4, 76.3]
- Random Forest Accuracy: [75.9, 75.6, 75.3, 75.7, 75.8]
- Differences: [0.6, 0.6, 0.5, 0.7, 0.5]

The test involved analyzing the variance in accuracy scores among the different models using the ANOVA (Analysis of Variance) test. ANOVA compares the mean accuracy scores of Decision Tree, Random Forest, AdaBoost, K-Nearest Neighbors, and Support Vector Machines to determine if the differences in performance are statistically significant. If the p-value obtained from the test was less than 0.05, it indicated that the observed variations in model accuracy were significant rather than due to random fluctuations. The results confirmed that the Decision Tree model's superior accuracy (76.31%) was statistically significant when compared to the other models, validating its effectiveness for predicting student academic performance.

To further clarify the relationship between SES factors and model performance, Table 5 presents the Statistically Significant test results for each model based on different SES factors.

Table 5
Presents the Statistical Significance test results for Models based on different SES factors

SES Factor	Decision Tree (%)	Random Forest (%)	AdaBoost (%)	KNN	SVM	F-Statistic	P-Value	interpretation
Study time	76.3	75.9	74.3	72.5	70.8	324.86	0.0	Significant
Aspirations of Higher Education	75.8	75.5	74.7	72.3	71.0	324.86	0.0	Significant
Family Dynamics	76.1	75.7	74.5	72.8	70.5	324.86	0.0	Significant
School Quality	76.5	75.6	74.2	72.4	71.2	324.86	0.0	Significant

The F-statistic measures the variance between the model performances, while the p-value determines whether the differences are statistically significant. If the p-value is less than 0.05, the SES factor significantly affects model performance. The interpretation column indicates whether each factor has a meaningful impact on model accuracy. Based on the results, the Decision Tree model consistently performed the best across all SES factors, achieving the highest accuracy scores in predicting student academic performance.

The Decision Tree performed the best with a 76.31% accuracy, which was statistically significant ($p < 0.05$) compared to the other models. Although deep learning models may provide higher accuracy, they lack transparency, making them less suitable for direct application in educational decision-making.

4.1 Interpretation and Discussion

The findings highlight several key insights:

i. Tree-based models (Decision Tree and Random Forest) performed best, emphasizing the importance of structured decision-making models in academic predictions.
ii. SES factors such as study time, aspirations, family dynamics, and school quality strongly influenced prediction accuracy, suggesting that educational policies should incorporate these variables for targeted interventions.
iii. Statistical validation using ANOVA confirmed that observed differences between model performances were significant, indicating that Decision Tree's superior accuracy was not due to chance.

5. Practical Implications for Educational Policy

While our study benchmarks machine learning models, it is equally important to extract actionable insights for educators and policymakers. The findings of this study provide a foundation for data-driven decision-making in education, enabling institutions to implement more effective student support strategies. The identification of key factors such as socioeconomic status, study habits, and parental involvement allows policymakers to develop targeted interventions that address disparities in academic performance. By leveraging interpretable machine learning models like Decision Trees, schools can monitor at-risk students in real-time and proactively implement personalized learning plans tailored to individual needs. Furthermore, educators can benefit from predictive analytics by receiving early warnings about students who may require additional academic support, enabling timely interventions before performance declines. Teacher training programs should also integrate machine learning insights to enhance instructional strategies, ensuring that teachers are equipped with the necessary tools to utilize predictive data effectively. Additionally, education ministries and institutions can use these findings to allocate resources more efficiently, directing support where it is most needed, such as tutoring programs, financial aid, or mentorship initiatives. The integration of predictive analytics into the educational system fosters a more equitable learning environment by reducing achievement gaps and ensuring that all students have access to the necessary support to succeed:

i. Early Intervention Programs: Institutions should monitor at-risk students based on SES factors and academic records.
ii. Personalized Learning Plans: Schools should use interpretable models like Decision Trees to create individualized student support strategies.
iii. Teacher Training: Educators should be trained on how to interpret and apply machine learning insights to enhance student engagement and performance.

6. Conclusion

This study provides a comparative analysis of classification algorithms for student academic performance prediction. While deep learning models are gaining popularity, this research highlights the enduring relevance of explainable machine learning models, particularly Decision Trees and Random Forests. The findings contribute to educational data mining by demonstrating that ML models can be used to develop actionable, transparent, and interpretable interventions.

A key aspect of this study is the emphasis on socioeconomic status (SES) factors, which play a crucial role in shaping student performance. The results reveal that SES-related variables such as parental education, family income, and access to academic resources significantly influence student achievement. By incorporating SES data into predictive models, this research underscores the need for targeted educational interventions that address disparities among students from different socioeconomic backgrounds. The study also provides insights into how educational institutions can leverage ML models to identify at-risk students early and tailor support strategies that enhance learning outcomes.

The findings of this research highlight the potential of machine learning in shaping educational policies, particularly in promoting equity and inclusivity in academic environments. Schools and policymakers can utilize these models to allocate resources more effectively, ensuring that students from lower SES backgrounds receive the necessary support to bridge academic gaps. The study also demonstrates that interpretable models like Decision Trees provide a transparent and practical approach for integrating machine learning into academic decision-making.

By integrating SES factors into predictive analytics, future research can refine these models to provide more personalized recommendations for student success. Future research should also explore hybrid models combining explainability with deep learning and validate findings using real-world student intervention trials.

Acknowledgement
This research was funded by a grant from Ministry of Higher Education of Malaysia (FRGS Grant R.J130000.7824.4X172).

References
[1] Farissi, Al, Halina Mohamed Dahlan, and Samsuryadi. "Genetic algorithm based feature selection for predicting student's academic performance." *In Emerging Trends in Intelligent Computing and Informatics: Data Science, Intelligent Information Systems and Smart Computing* 4, pp. 110-117. Springer International Publishing, 2020. https://doi.org/10.1007/978-3-030-33582-3_11
[2] Rodríguez-Hernández, Carlos Felipe, Eduardo Cascallar, and Eva Kyndt. "Socio-economic status and academic performance in higher education: A systematic review." *Educational Research Review* 29 (2020): 100305. https://doi.org/10.1016/j.edurev.2019.100305
[3] Lutfiu, Besim, and Laura Lutfiu Hoxha. "Socio-Economic Status of Students and Its Impact on the Quality of Studies." *European Journal of Education and Pedagogy* 5, no. 4 (2024): 66-72. https://doi.org/10.24018/ejedu.2024.5.4.852
[4] Huang, Y., & Chen, X. Predicting academic achievement using demographic and behavioral variables. Journal of Applied Educational Research 45 no. 1 (2017): 78-95.
[5] Shyama Heshini Niranjala, Mustafa Muwafak Alobaedy, and S B Goyal. "A Comparative Study of Machine Learning Techniques for Predicting Student Academic Performance." *Lecture Notes in Networks and Systems* (2024): 307–15. https://doi.org/10.1007/978-3-031-73318-5_31
[6] Rabia, Mahwish, Naima Mubarak, Hira Tallat, and Wajiha Nasir. "A study on study habits and academic performance of students." *International Journal of Asian Social Science* 7, no. 10 (2017): 891-897. http://dx.doi.org/10.18488/journal.1.2017.710.891.897
[7] H. Alamri, Leena, Ranim S. Almuslim, Mona S. Alotibi, Dana K. Alkadi, Irfan Ullah Khan, and Nida Aslam. "Predicting student academic performance using support vector machine and random forest." In *Proceedings of the 2020 3rd international conference on education technology management*, pp. 100-107. 2020. https://doi.org/10.1145/3446590.3446607
[8] Wen, Xiao, and Hu Juan. "Early prediction of students' performance using a deep neural network based on online learning activity sequence." *Applied Sciences* 13, no. 15 (2023): 8933. https://doi.org/10.3390/app13158933
[9] Chen, Sa, and Xiankun Lin. "Application of Decision Tree Algorithm in Educational Data Mining." *Curriculum and Teaching Methodology* 6, no. 8 (2023): 120-127. https://doi.org/10.23977/curtm.2023.060818
[10] Munir, Javeria, Mehreen Faiza, Bakht Jamal, Sana Daud, and Khurshed Iqbal. "The impact of socio-economic status on academic achievement." *Journal of Social Sciences Review* 3, no. 2 (2023): 695-705. http://dx.doi.org/10.54183/jssr.v3i2.308
[11] Li, X., & Chen, Y. "Integrating Support Vector Machines and Decision Trees for Academic Achievement Analysis" (2019).
[12] Hamsa, Hashmia, Simi Indiradevi, and Jubilant J. Kizhakkethottam. "Student academic performance prediction model using decision tree and fuzzy genetic algorithm." *Procedia technology* 25 (2016): 326-332.
[13] Cortez, P. *Student Performance [Dataset]. UCI Machine Learning Repository*. 2014. https://doi.org/10.1016/j.protcy.2016.08.114