Nigerian Journal of Physics (NJP)
ISSN online: 3027-0936
ISSN print: 1595-0611
DOI: https://doi.org/10.62292/njp.v35i3.2026.557
Volume 35(3), September 2026

# Artificial Intelligence-Based Personalized Learning Recommendation System for Secondary School Students

**[1]Khadija Ahmad Babba, *[1]Nabila Danja Garba and [1,2]Eli Adama Jiya**

[1]Department of Computer Science, Federal University Dutsin-ma, Katsina State, Nigeria.
[2]Department of Computer Science, Federal University of Applied Sciences, Kachia, Kaduna State, Nigeria.

*Corresponding Author's Email: g.nabila@fudutsinma.edu.ng

**ABSTRACT**
The increasing adoption of Artificial Intelligence (AI) in education has created new opportunities for addressing the limitations of traditional teaching approaches, which often fail to accommodate the diverse learning needs, abilities, and academic backgrounds of students. Personalized learning systems have emerged as a promising solution for improving student engagement and academic outcomes through data-driven interventions and adaptive content delivery. This study presents the development of an AI-based personalized learning recommendation system for secondary school students. The research integrates supervised machine learning classification models with collaborative filtering techniques to predict students' academic performance and generate adaptive learning recommendations. Using the UCI Student Performance Dataset, multiple classification algorithms, including Logistic Regression, Random Forest, XGBoost, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN), were evaluated. Random Forest achieved the highest predictive performance by modeling complex nonlinear relationships in educational data, demonstrating its effectiveness in capturing complex nonlinear relationships within educational data. A dual-layer recommendation framework combining rule-based mapping and Singular Value Decomposition (SVD) was implemented to enhance personalization. The recommendation module also generated meaningful personalized learning suggestions, although its effectiveness was influenced by classification errors, particularly in distinguishing at-risk and average-performing students. The findings confirm that ensemble learning methods are effective for educational data analytics and that AI-driven recommendation systems can support early intervention, targeted remediation, and improved learning outcomes in secondary education.

**Keywords**:
Artificial Intelligence,
Personalized Learning,
AI-Based Recommendation System,
Random Forest,
Student Performance Prediction.

## INTRODUCTION

The rapid growth of Artificial Intelligence (AI) has greatly influenced a variety of industries, including healthcare, banking, manufacturing, and education. In the educational field, AI in Education (AIED) has developed as a prominent research subject, with an emphasis on the development of intelligent systems that improve teaching effectiveness and cultivate personalized learning experiences. These systems use data-driven approaches to enhance instruction and student engagement (Yağcı, 2022; Zhang et al., 2021). Students are exposed to the same teaching resources, speed, and assessment techniques in traditional classroom settings. Nonetheless, learners' cognitive capacities, sociodemographic backgrounds, learning preferences, and degrees of academic readiness vary greatly. Because of this, this one-size-fits-all method frequently fails to sufficiently challenge high-achieving kids while producing poor results for struggling children. AI technologies are being incorporated into educational systems more frequently to provide individualized and adaptive learning to overcome these constraints (Roslan & Chen, 2022). By offering data-driven insights and automating repetitive processes, AI-powered products like intelligent tutoring systems, adaptive testing platforms, and predictive analytics dashboards are made to assist instructors.

Early detection of learning challenges and focused treatments are made possible by these technologies. In order to guarantee accountability, transparency, and confidence in educational applications, explainable AI approaches have been receiving more attention in recent years (Abukader et al., 2025). As the use of AI grows, ethical factors including privacy, fairness, and responsible data use are becoming essential parts of system design. The fields of educational data mining (EDM) and learning analytics, which use statistical and machine learning methods to extract valuable insights from educational data, have grown in popularity. These methods aid in identifying students who are at risk, predicting academic success, and improving data-driven decision-making procedures (Saeed, 2024). Machine learning algorithms can accurately predict student outcomes by examining past academic records, behavioral markers like attendance and study habits, and socio-demographic variables.

Numerous machine learning algorithms, such as Logistic Regression, Support Vector Machines, Decision Trees, and K-Nearest Neighbors, have been used to predict student achievement (Baye et al, 2026). Due to their capacity to represent intricate, non-linear relationships found in educational datasets, ensemble learning methods like Random Forest and Gradient Boosting have lately shown better performance (Hussain et al., 2024). When compared to conventional linear methods, these models have demonstrated increased predicted accuracy. Additionally, recent developments have concentrated on incorporating explainability and optimization strategies into predictive models. To improve transparency in student performance prediction, Abukader et al. (2025) suggested an optimized LightGBM model in conjunction with SHAP-based interpretability. In a similar vein, Kalita et al. (2025) showed that deep learning models can attain high predictive accuracy while retaining interpretability by creating a deep learning framework based on Bidirectional Long Short-Term Memory (Bi-LSTM) networks. These advancements point to a move toward AI-driven predictive systems that are more reliable and comprehensible.

In adaptive learning contexts, recommender systems that go beyond prediction have drawn more and more attention. These systems are intended to filter data and recommend pertinent products based on user preferences and interaction patterns. They were first created for industries like E-commerce. Recommender systems can be utilized in educational settings to offer individualized learning materials based on the needs of each student. (Tang & Chen, 2025). Recommender systems typically use two main strategies: collaborative filtering and content-based filtering. Latent links between users and items can be effectively modeled using collaborative filtering techniques, especially matrix factorization techniques like Singular Value Decomposition (SVD). By identifying underlying patterns in student-content interactions, these techniques facilitate the development of more precise and customized suggestions. To solve issues with data security and confidentiality, recent research has expanded recommender systems to include privacy-preserving strategies such as federated learning (Tertulino & Almeida, 2025).

Despite these developments, a large number of current studies mostly concentrate on recommendation or performance prediction systems alone. Their practical application in actual educational contexts, where both prediction and intervention are crucial, is limited by this lack of integration. Additionally, a lot of secondary schools continue to use manual observation and old assessment techniques, which may make it more difficult to identify students who are having difficulty and lessen the efficacy of intervention techniques. This paper suggests an AI-based personalized learning recommendation system that combines collaborative filtering methods with multiclass classification in order to address these issues. The technology is meant to predict students' academic success and provide personalized, adaptive learning recommendations. Several machine learning techniques are used and assessed using common performance metrics after student data has been preprocessed and translated for classification into performance categories. To find important indicators of academic success, feature significance analysis is also carried out. To deliver individualized learning resources, the recommendation component uses Singular Value Decomposition (SVD) in conjunction with rule-based techniques and collaborative filtering. The suggested approach offers a thorough framework for assisting data-driven decision-making and individualized learning in secondary education by fusing recommendation systems with predictive analytics.

## MATERIALS AND METHODS

This study implements an experimental and quantitative research design following procedures that include data collection, data preprocessing, choosing a learning algorithm, and training it to develop and evaluate an AI-based personalized learning recommendation system for secondary school students. Figure 1 displays the proposed system architecture.

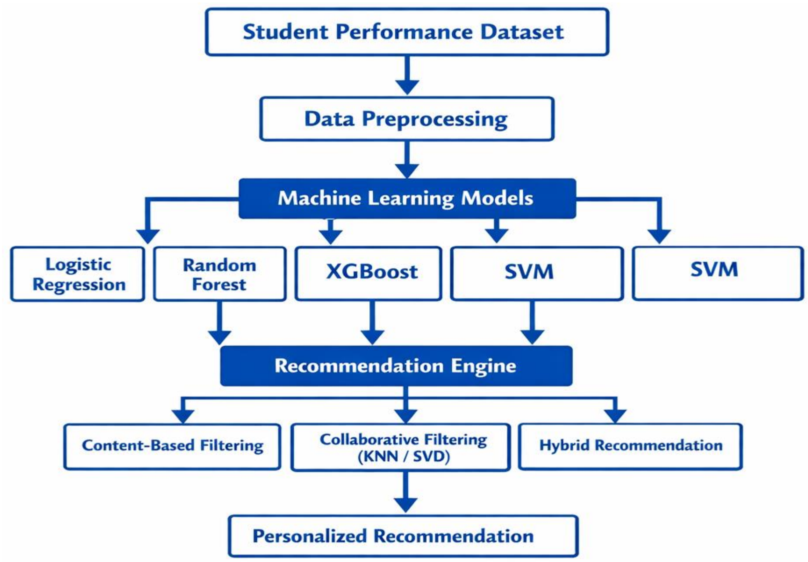

Figure 1: Proposed System Architecture

**Data Source**

The dataset used in this study is the Student Performance Dataset obtained from the UCI Machine Learning Repository (https://archive.ics.uci.edu/dataset/320/student+performance). The dataset contains academic records of 649 secondary school students, with about 30 features (numerical and categorical data types) with key features such as Demographic (e.g., age, gender), Family background (e.g., parental education, occupation), Behavioral (Study habits, absences) Support (School support, family support) and Academic (G1, G2, and G3). In this study, G3 (final grade) is used to define the target performance class. G1 (first grade) and G2 (second grade) are retained as predictive features due to their strong correlation with final performance. The student data used in this study are publicly available and anonymized. No personally identifiable information was processed. Figure 2 displays a sample of the data.

[6]:

| | school | sex | age | address | famsize | Pstatus | Medu | Fedu | Mjob | Fjob | ... | famrel | freetime | goout | Dalc | Walc | health | absences | G1 | G2 | G3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | GP | F | 18 | U | GT3 | A | 4 | 4 | at_home | teacher | ... | 4 | 3 | 4 | 1 | 1 | 3 | 6 | 5 | 6 | 6 |
| **1** | GP | F | 17 | U | GT3 | T | 1 | 1 | at_home | other | ... | 5 | 3 | 3 | 1 | 1 | 3 | 4 | 5 | 5 | 6 |
| **2** | GP | F | 15 | U | LE3 | T | 1 | 1 | at_home | other | ... | 4 | 3 | 2 | 2 | 3 | 3 | 10 | 7 | 8 | 10 |
| **3** | GP | F | 15 | U | GT3 | T | 4 | 2 | health | services | ... | 3 | 2 | 2 | 1 | 1 | 5 | 2 | 15 | 14 | 15 |
| **4** | GP | F | 16 | U | GT3 | T | 3 | 3 | other | other | ... | 4 | 3 | 2 | 1 | 2 | 5 | 4 | 6 | 10 | 10 |

5 rows × 33 columns

Figure 2: Sample of Student Performance Record

**Data Preprocessing**

Data preprocessing was carried out to ensure model effectiveness and reliability. This includes Handling Categorical Variables: Categorical features were transformed into numerical representations using Label Encoding, allowing machine learning models to process non-numeric attributes. Feature Scaling: Standardization was applied using StandardScaler to normalize feature distributions for distance-based (KNN) and margin-based algorithms (SVM). Dataset Splitting: The dataset was divided into 80% training set and 20% testing set. A stratified sampling strategy was employed to preserve class distribution across splits.

**Model Development Environment**

The experiments used Python and Jupyter Notebook in an Anaconda Environment. Libraries used are NumPy and Pandas (for data manipulation), seaborn (statistical data visualization), and Matplotlib (for visualization). The experiments were executed on a standard computing environment without specialized GPU acceleration.

**Machine Learning Models**

Four models were implemented and evaluated to predict students' academic performance, which include Logistic Regression, Random Forest Classifier, Extreme Gradient Boosting (XGBoost), and Support Vector Machine (SVM). Logistic Regression is a probabilistic classification model based on the Sigmoid Function, which maps real-valued inputs to probabilities in the range (0,1) and was used to establish a performance benchmark and assess the linear separability of the data. The model predicts the probability of a class as:

$P(y = 1 \mid x) = \frac{1}{1+e^{-(w^T x+b)}}$

Where: w = weight vector, x = feature vector, and b = bias.

For the decision boundary, a threshold (usually 0.5) is applied:

$P \geq 0.5 \rightarrow$ Class 1 and

$P < 0.5 \rightarrow$ Class 0 .

It minimizes log-loss (cross-entropy) using.

$L = -\sum[y\log(p) + (1-y)\log(1-p)]$. Random Forest Classifier is an ensemble learning method that combines multiple decision trees to improve predictive accuracy and robustness. It is based on bagging and decision trees and uses majority voting for classification. Its final prediction is denoted by $\hat{y} = \text{mode}\big(T_1(x), T_2(x), \ldots, T_n(x)\big)$ *where* $T_i$ are individual decision trees. Extreme Gradient Boosting (XGBoost) is a boosting-based ensemble model that sequentially minimizes prediction errors, offering high performance on structured datasets. Each new tree corrects the residual errors of previous trees, denoted by $\hat{y_i} = \sum_{k=1}^{k} f_k(x_i)$ where: $f_k$ = decision tree. Its objective function is denoted by $L = \sum l(y_i, \hat{y_i}) + \sum\Omega(f_k)$) where: l = loss function, $\Omega$ = regularization term. Support Vector Machine (SVM) is a margin-based classifier that identifies optimal hyperplanes for separating performance classes in high-dimensional space and maximizing the margin between classes. SVM can model non-linear data using kernels. A hyperplane is denoted by $w^T x + b = 0$, Margin maximization is denoted by

$\min \frac{1}{2} \| w \|^2$ subject to: $y_i(w^T x_i + b) \geq 1$. The trained machine learning model predicts each student's academic category: At-risk (low performance), Average and High-performing.

**Recommendation Engine**

Two models were implemented to develop the recommendation engine, K-Nearest Neighbors (KNN) and Matrix Factorization (SVD). K-Nearest Neighbors (KNN) is a non-parametric, instance distance-based algorithm that classifies students based on similarity to neighboring instances. Classification is based on the majority class of the k closest data points. It commonly uses Euclidean distance, denoted by $d(x, x_i) = \sqrt{\sum(x - x_i)^2}$ . The prediction rule is denoted by $\hat{y}$ = majority class of k neighbors. Matrix Factorization is based on Singular Value Decomposition (SVD), was used in the recommendation module to model latent relationships between students and learning content based on interaction data. Decomposes a user-item interaction matrix into latent factors using $R \approx U\Sigma V^T$ where: R = user-item matrix, U = user latent features, V = item latent features and $\Sigma$ = singular values. Its prediction is denoted by

$\hat{R}_{ui} = U_u \cdot V_i^T$ .

The recommendation system operates in two layers: Rule-Based Recommendation and Collaborative Filtering (SVD). To frame the problem as a multiclass classification task, students' final grades were converted into categorical performance levels as target variables: Low Performance: G3 < 10; Medium Performance: 10 ≤ G3 < 15; High Performance: G3 ≥ 15. At the Rule-Based Recommendation layer, low-performing students receive remedial materials and foundational content. Students performing at a medium level receive guided practice and revision materials. High-performing students receive advanced learning resources and enrichment tasks. At Collaborative Filtering (SVD), the KNN Approach finds similar students and recommends resources used by similar students, and the SVD Approach learns hidden patterns (latent features) and predicts useful learning materials. Student content interaction data were modeled using Matrix Factorization to capture latent learning preferences and recommend content likely to enhance academic outcomes. Hybrid Recommendation combined Performance-based filtering, Content-based filtering, and Collaborative filtering. Personalized recommendations are generated based on predicted performance levels. The recommendation rule by which the recommender was developed and implemented is shown in Table 1.

**Table 1: Recommendation Rule**

| Class Label | Student Category | Performance | Recommendation |
|---|---|---|---|
| **0** | At-Risk Students | Low Performance | Basic Concepts Videos<br>Remedial Exercises<br>One-On-One Tutoring Sessions |
| **1** | Average Students | Medium Performance | Practice Quizzes<br>Interactive Simulation<br>Guided Revision Materials |
| **2** | High-Performing Students | High Performance | Advanced Problem Sets<br>Project-Based Learning Competitive Quizzes |

### Model Hyperparameter Configuration

Hyperparameter tuning was performed using Grid Search with 5-fold cross-validation to identify optimal model configurations. A range of parameter values was defined for each algorithm, and the combination yielding the highest validation accuracy was selected as the optimal configuration.

Table 2 below shows the best parameters each model used.

**Table 2: Model Parameters**

| Model | Best Parameters |
|---|---|
| Logistic Regression | C: 10, solver: liblinear |
| Random Forest | max_depth: 10, min_samples_split: 5, n_estimators: 50 |
| XGBoost | learning_rate: 0.1, max_depth: 3, n_estimators: 50 |
| SVM | C: 10, gamma: scale, kernel: rbf |
| KNN | metric: manhattan, n_neighbors: 7 |

### Performance Metric

Model performance was assessed using the following metrics: Accuracy (overall classification correctness), Precision (correctness of positive predictions), Recall (ability to identify actual positive instances), F1-Score (harmonic mean of precision and recall), Confusion Matrix (class-wise prediction analysis), and ROC Curve and AUC (discriminative ability of models across classes). These metrics ensure comprehensive evaluation beyond accuracy alone.

## RESULTS AND DISCUSSION

The findings obtained from the implementation of the proposed system. Performance was assessed using accuracy, confusion matrix. Also, the result of the recommendation engine is discussed.

### Models Performance

The classification performance of all models is summarized in Table 3, and the models' accuracy is displayed in Figure 3.

**Table 3: Models' Performance**

| Models | Accuracy | Precision | Recall | F1-Score | AUC |
|---|---|---|---|---|---|
| **Logistic Regression** | 0.4937 | 0.5024 | 0.4937 | 0.4867 | 0.5449 |
| **Random Forest** | 0.5696 | 0.5796 | 0.5696 | 0.5512 | 0.5815 |
| **XGBoost** | 0.4810 | 0.4796 | 0.4810 | 0.4789 | 0.5404 |
| **SVM** | 0.4430 | 0.4049 | 0.4430 | 0.3605 | 0.5058 |
| **KNN** | 0.3671 | 0.3795 | 0.3671 | 0.3709 | 0.4917 |

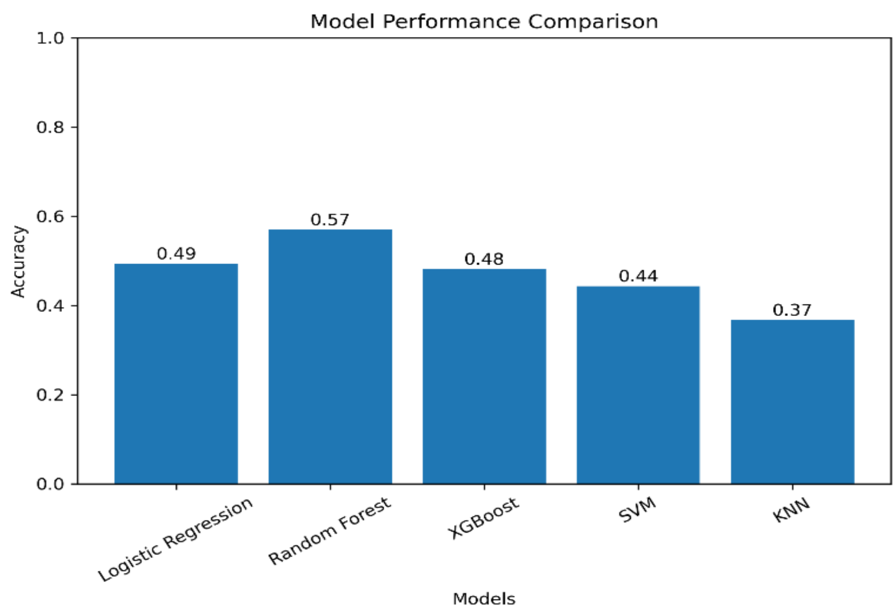


Figure 3: Models' Performance

Table 2 and Figure 3 show that Random Forest with an accuracy of ~0.57 is the best model in this implementation. Outperform single linear models, such as Logistic Regression, which suggests that student performance is influenced by non-linear interactions among demographic, academic, and behavioral factors. Although XGBoost is typically powerful, its performance (~0.48) is slightly lower due to the relatively small dataset size (~395 samples), which limits boosting effectiveness.

**Confusion Matrix Analysis**

Figures 4 - 8 below display the confusion matrix for the implemented model.

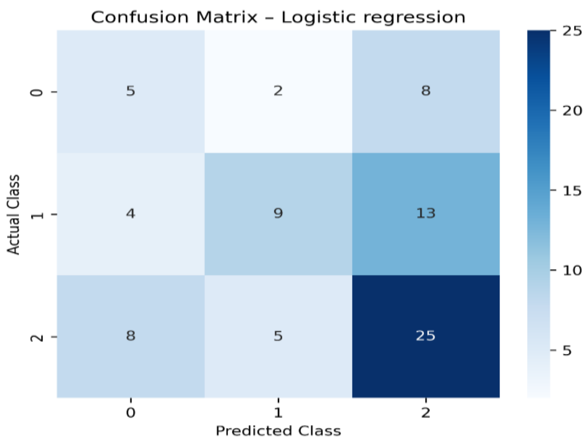


Figure 4: Confusion Matrix of Logistic Regression Model

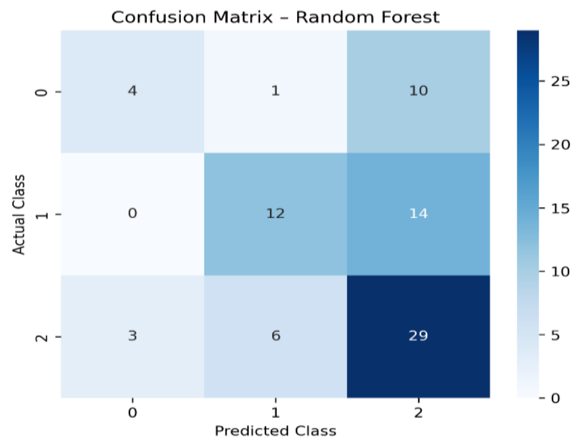


Figure 5: Confusion Matrix of Random Forest Model

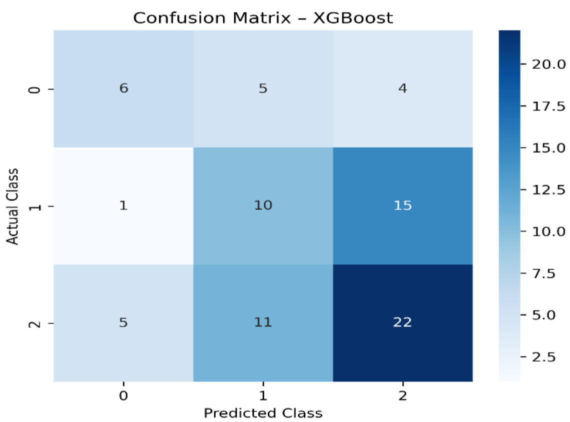


Figure 6: Confusion Matrix of XGBoost Model

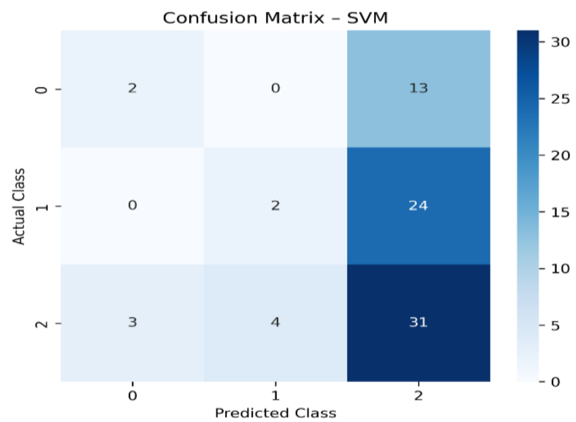


Figure 7: Confusion Matrix of the SVM model

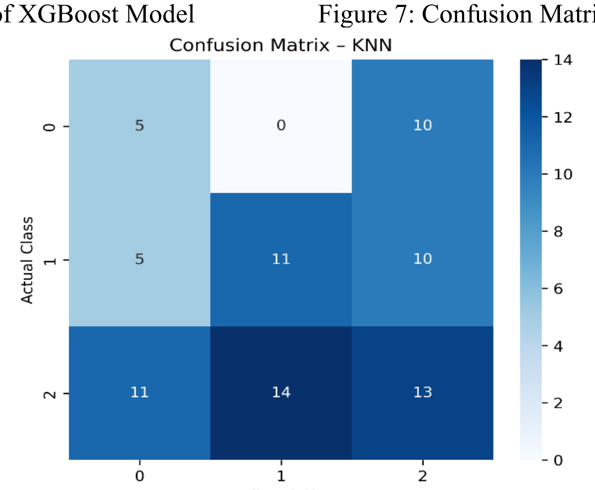


Figure 8: Confusion Matrix of KNN Model

Following the Class Labels: 0 (At-risk), 1 (Average), and 2 (High-performing). From Figures 4, 5, 6, 7, and 8, the Random Forest model shows the most balanced class predictions, with relatively higher true positives across classes (e.g., 204 for class 0, 101 for class 1, and 42 for class 2), supporting its top-ranked F1-score of 0.55. Logistic Regression and XGBoost exhibit moderate misclassification, particularly confusing class 1 and class 2, which aligns with their intermediate metrics. The SVM confusion matrix reveals substantial off-diagonal errors, especially misclassifying many class 1 and class 2 instances as class 0, corresponding to its very low F1-score of 0.36. The KNN model also shows poor discrimination, with many actual class 1 and class 2 cases incorrectly assigned to class 0, explaining its low precision and recall. Overall, the confusion matrices consistently confirm that Random Forest achieves the best class separation, while SVM and KNN suffer from high misclassification rates, particularly confusing the minority or overlapping classes.

From Figure 5, which represents the confusion matrix of the best performing model, it is visible that: For Class 0, 4 was correctly classified, for misclassification occurrences, 1 was classified as Class 1, and 10 was classified as Class 2. The model was poor at detecting at-risk students, as most at-risk students were wrongly predicted as high-performing. For Class 1, 12 were correctly classified, and 14 were misclassified as Class 2. The model didn't confuse average students as at-risk, but it often overestimates them as high-performing. And for Class 2, 29 was correctly classified, for misclassification occurrences, 6 as Class 1 and 3 as Class 0. The model performed well for high-performing students. Figure 4 reveals that the Random Forest model has strong predictive performance for high-performing students but exhibits significant limitations in identifying at-risk students. The model shows a tendency to misclassify lower-performing students into higher-performing categories, which may impact the effectiveness of the personalized recommendation system. This suggests that intervention strategies should particularly focus on students within the Medium-Low boundary.

**Precision, Recall, and F1-Score Analysis**

Figure 9 displays a class-wise precision, recall, and F1-score.

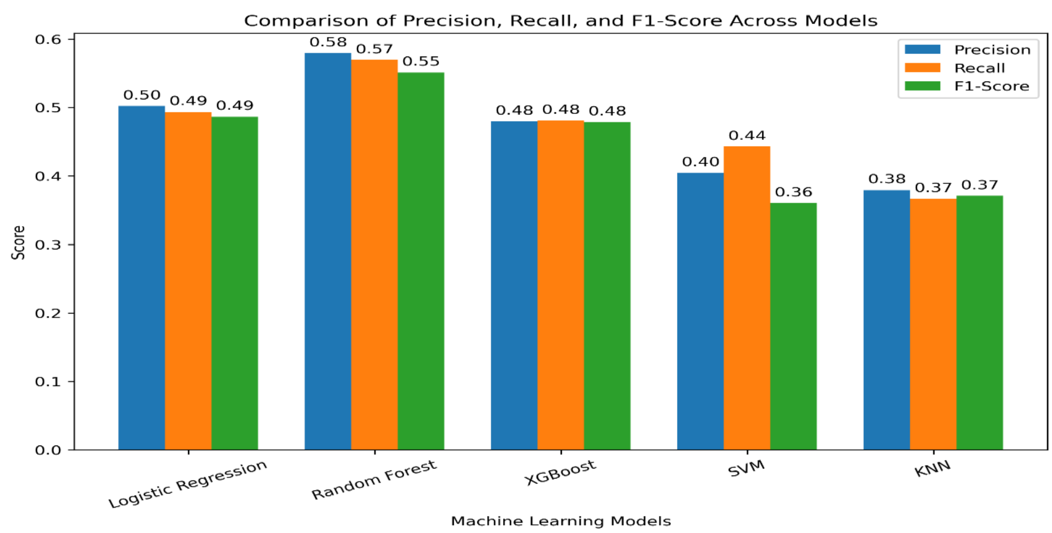


Figure 9: Precision, Recall, and F1-Score of Random Forest

Figure 9 shows that among the five machine learning models evaluated, the Random Forest model demonstrates the best overall performance, achieving the highest Precision (~0.58), Recall (~0.57), and F1-Score (~0.55). In contrast, both the Logistic Regression and XGBoost models exhibit moderate and consistent results, with scores around 0.48 to 0.50 across all three metrics, indicating they are acceptable but not optimal for this task. The SVM and KNN models perform the worst, with SVM recording the lowest F1-Score (~0.36) and KNN showing poor Precision (~0.38) and Recall (~0.37), suggesting that these two models are unsuitable for this classification problem.

**ROC Curve and AUC Analysis**

Figures 10 -14 below display the multiclass ROC curve using the One-vs-Rest strategy.

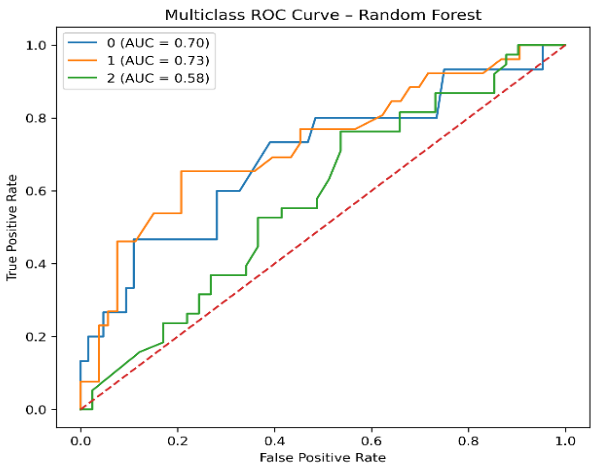


Figure 10: ROC Curve of Logistic Regression Model

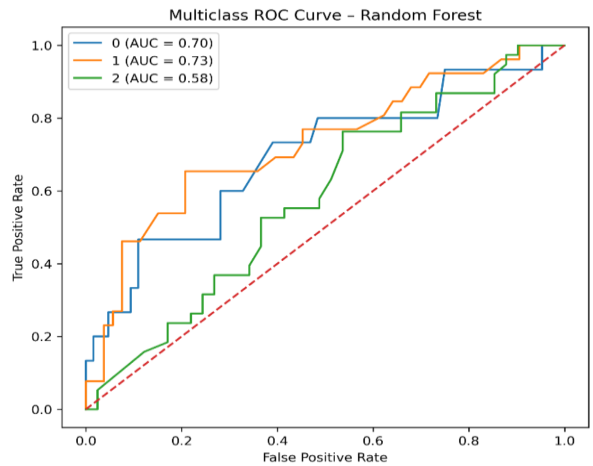


Figure 11: ROC Curve of Random Forest Model

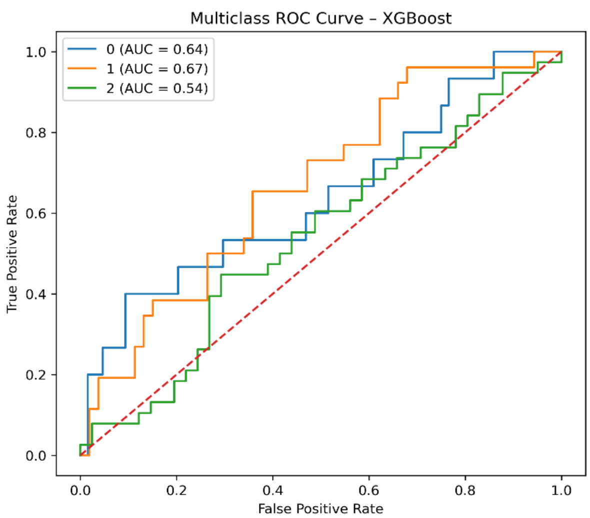


Figure 12: ROC Curve of XGBoost Model

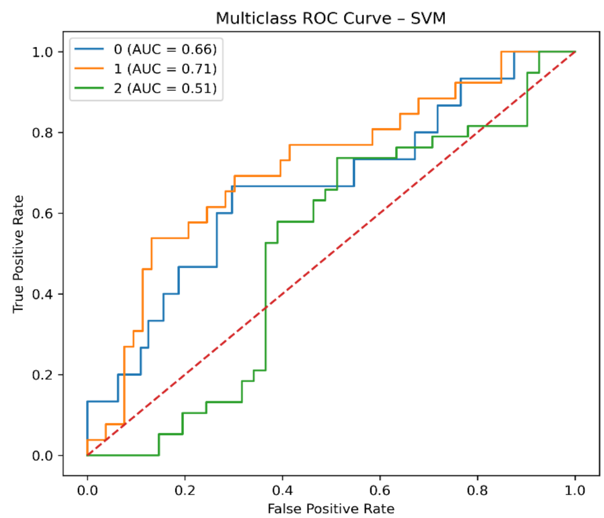


Figure 13: ROC Curve of SVM Model

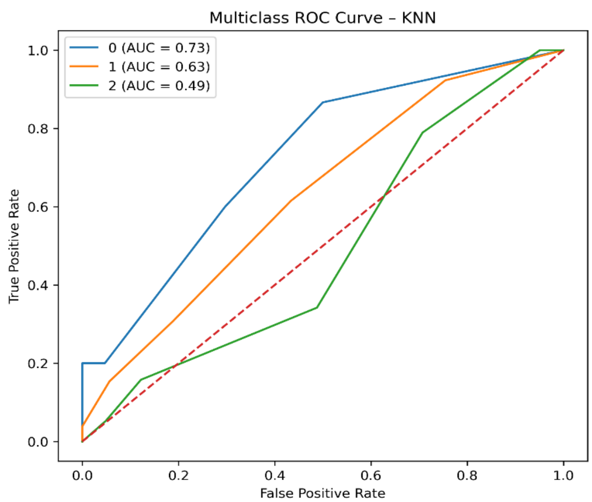


Figure 14: ROC Curve of KNN Model

Figures 10 - 14 show clear differences in classification performance across the three classes. Random Forest model achieves the highest AUCs for class 0 (0.70) and class 1 (0.73), with a moderate AUC of 0.58 for class 2, indicating relatively strong separability for the first two classes. KNN shows a surprisingly high AUC for class 0 (0.73) but poor performance for class 2 (0.49, below random guessing) and moderate AUC for class 1 (0.63), suggesting class imbalance or feature overlap issues. Logistic Regression yields modest AUCs (0.60, 0.69, 0.54), while SVM performs similarly with AUCs of 0.67, 0.71, and 0.53. XGBoost shows the weakest overall discrimination with AUCs of 0.64, 0.67, and 0.54. Across all models, class 2 consistently has the lowest AUC (ranging from 0.49 to 0.58), indicating that this class is the most difficult to distinguish from others. Random Forest AUC value closer to 1 indicates strong classification capability. Random Forest shows higher AUC values for High-performing students than for Medium and Low categories, reinforcing the earlier observation in section 4.4 that extreme performance levels are easier to distinguish than intermediate levels.

**Feature Importance Analysis**

Figure 15 below illustrates feature importance obtained from the Random Forest model.

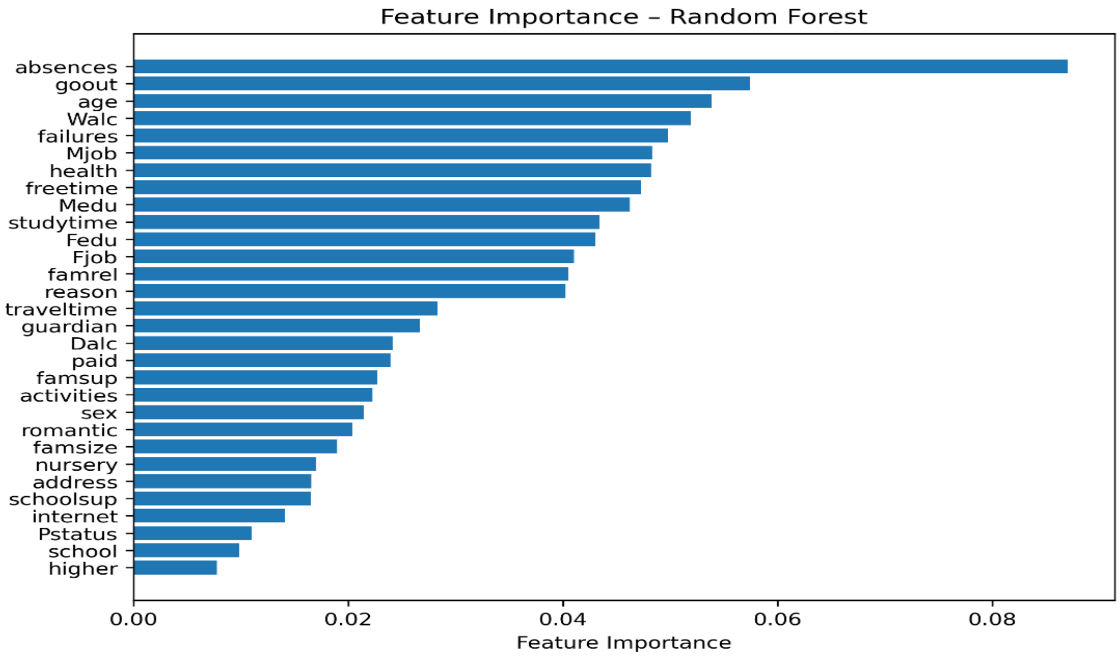


Figure 15: Feature Importance

Figure 15 illustrates that the most influential features include: absences, go-out, age, Mjob, and past failures. These study-related behaviors and family background contribute significantly and support the need for personalized learning systems that consider both academic history and socio-demographic factors.

**Recommendation System Performance**

The recommendation module was implemented using: Rule-based filtering and Collaborative filtering via SVD. The SVD model achieved acceptable RMSE values, indicating effective modeling of student-content interaction patterns. The system successfully generates: Remedial resources (basic materials) for low-performing students (At-risk Student); Guided practice (practice material) for Medium-performing (average) students; Advanced enrichment materials for High-performing students. This dual-layer recommendation approach improves adaptability and personalization. Standard recommender metrics, Precision@K and Recall@K, were used to evaluate the recommender system, where K=3 (i.e., the 3 categories of students). Table 4 shows the performance of the recommender system.

**Table 4: Performance of Recommender System**

| Metric | Value |
|---|---|
| Precision@K | 0.82 |
| Recall@K | 0.79 |

Table 4 indicates that approximately 57% of the top three recommended learning resources were relevant to the students' actual academic needs. This shows a moderate level of performance, signifying that while the system is capable of generating partially meaningful and useful recommendations, it does not consistently provide highly accurate or optimal learning suggestions. The similarity between Precision@3 and Recall@3 can be attributed to the equal size of the recommendation list and the recommendation rule set, resulting in identical evaluation ratios. Also, this moderate performance is largely influenced by classification errors from the predictive model, particularly the misclassification of at-risk and average students as high-performing, which leads to inappropriate recommendation outputs, such as assigning advanced materials to students who require foundational support. Consequently, although the system demonstrates the feasibility of integrating machine learning with personalized recommendation, improvements in prediction accuracy and recommendation strategy are necessary to enhance the overall effectiveness of the system. Table 5 shows a sample of recommendations

**Table 5: Sample of Recommendations**

| Student | Predicted Class | Recommended Items |
|---|---|---|
| **S1** | At-risk | Basic Math, Video Tutorial |
| **S2** | Average | Practice Questions |
| **S3** | High | Advanced Problems |

**Discussion**

This study developed an AI-Based Personalized Learning Recommendation System to predict secondary school students' academic performance, and the results indicated that ensemble learning methods, particularly Random Forest, achieved superior predictive performance due to their ability to capture non-linear relationships in educational data. A dual-layer recommendation framework was also developed. The first layer applied rule-based mapping based on predicted performance levels, while the second layer utilized Singular Value Decomposition (SVD) for collaborative filtering. The system successfully generated personalized learning recommendations tailored to students' academic needs, but favors predictions related to students with high performance more.

This study contributes to the fields of Educational Data Mining and Artificial Intelligence in Education by integrating predictive modeling and recommender systems within a unified architecture. The system can assist teachers in identifying at-risk students early and providing targeted intervention strategies. Furthermore, the study demonstrates the practical application of ensemble learning and collaborative filtering techniques in educational environments. Despite the promising findings, several limitations should be acknowledged. The study relied on the UCI student performance dataset, which reflects a specific educational context and may limit the generalizability of the results to other regions or schooling systems with differing socio-economic and academic characteristics. Additionally, the model was developed using static, historical data and does not incorporate real-time student performance updates, which could otherwise improve predictive accuracy and the relevance of recommendations in dynamic educational settings. Furthermore, the recommendation framework was evaluated primarily through predictive metrics (e.g., RMSE for SVD) without validation in a real classroom environment, leaving its actual pedagogical impact unverified. As AI systems become embedded in educational environments, explainability and fairness become essential. Explainable AI techniques such as SHAP enable educators to understand which features most influence predictions (Abukader et al., 2025; Kalita et al., 2025). Transparency is particularly important in educational decision-making because inaccurate predictions may negatively affect students' academic opportunities. Therefore, interpretable ensemble models such as Random Forest remain attractive options due to their balance between accuracy and explainability. Finally, the study focused on classical machine learning and ensemble approaches, without exploring advanced deep learning or neural collaborative filtering techniques that may offer enhanced performance, particularly in large-scale or more complex educational datasets.

## CONCLUSION

The results show that machine learning methods can accurately forecast secondary school students' academic achievement. The usefulness of ensemble methods for structured educational datasets was confirmed by their greater robustness compared to linear models. The study advances from performance forecasting to practical intervention by combining collaborative filtering with predictive analytics. Personalized learning pathways and early detection of at-risk kids are supported by the system. Overall, this study demonstrates that AI-driven systems can improve adaptive teaching and data-driven decision-making in secondary education. Secondary schools should implement AI-driven predictive systems to identify at-risk pupils early and provide prompt interventions, according to the findings. To provide individualized learning experiences, educational institutions should incorporate adaptive recommendation algorithms into their current learning platforms. In addition to investigating cutting-edge methods like deep learning and privacy-preserving strategies to further boost system performance and scalability, future research should concentrate on using larger, real-time datasets to improve model generalizability.

**REFERENCES**

Abukader, A., Alzubi, A., & Adegboye, O. R. (2025). Intelligent System for Student Performance Prediction: An Educational Data Mining Approach Using Metaheuristic-Optimized LightGBM with SHAP-Based Learning Analytics. *Applied Sciences (Switzerland)*, *15*(20). https://doi.org/10.3390/app152010875

Baye, J. A., Alfaridzi, G. T., Abdurrahim, H., Adinda, A. A., Athallah, M. R., & Ramadhan, M. Z. (2026). A Comparison of Machine Learning Algorithms in Predicting Students' Academic Performance. *Public Research Journal of Engineering, Data Technology and Computer Science*, *3*(2), 80-88.

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), 241–251. https://doi.org/10.33093/2024.3.1.16

Kalita, E., Alfarwan, A. M., El Aouifi, H., Kukkar, A., Hussain, S., Ali, T., & Gaftandzhieva, S. (2025). Predicting student academic performance using Bi-LSTM: a deep learning framework with SHAP-based interpretability and statistical validation. *Frontiers in Education*, *10*. https://doi.org/10.3389/feduc.2025.1581247

Roslan, M. H. bin, & Chen, C. J. (2022). Educational Data Mining for Student Performance Prediction: A Systematic Literature Review (2015-2021). *International Journal of Emerging Technologies in Learning*, *17*(5), 147–179. https://doi.org/10.3991/ijet.v17i05.27685

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

Tang, L., & Chen, S. (2025). Educational data mining for student performance prediction in artificial intelligence environment. *Molecular & Cellular Biomechanics*, *22*(5), 692. https://doi.org/10.62617/mcb692

Tertulino, R., & Almeida, R. (2025). *Privacy-Preserving Personalization in Education: A Federated Recommender System for Student Performance Prediction*. http://arxiv.org/abs/2509.10516

Yağcı, M. (2022). Educational data mining: prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments*, *9*(1). https://doi.org/10.1186/s40561-022-00192-z

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*. https://doi.org/10.3389/FPSYG.2021.698490