IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429 Volume 6, Issue 1, June 2026 Impact Factor: 8.2

Review of Hybrid Admixture Systems on the Strength and Durability Characteristics of Self-Compacting Concrete

Ram Jaysing Salunke[1] and Dr. Rajeev Kumar[2]

[1]Research Scholar, Department of Mathematics

[2]Associate Professor, Department of Mathematics

Vikrant University Gwalior M.P

Abstract: *Machine learning (ML) has emerged as an important analytical approach for predicting and understanding students' educational performance. The increasing availability of academic, demographic, behavioural, attendance, assessment, and learning-management-system data has created opportunities to identify factors associated with student achievement and to predict academic outcomes at an early stage. This review examines the major machine learning algorithms used for educational performance prediction, including Decision Tree, Random Forest, Support Vector Machine, k-Nearest Neighbour, Logistic Regression, Naïve Bayes, Artificial Neural Networks, Gradient Boosting, and ensemble-based approaches. The review focuses on their methodological characteristics, predictive capabilities, strengths, limitations, and suitability for different educational datasets. Traditional statistical approaches can provide useful interpretations, but machine learning methods are capable of modelling complex and nonlinear relationships among multiple educational variables.*

Tree-based and ensemble algorithms are particularly useful when datasets contain heterogeneous variables, while neural-network approaches can provide strong predictive performance when sufficient training data are available. However, issues such as data quality, class imbalance, missing values, overfitting, interpretability, privacy, and algorithmic bias remain important challenges. The review further highlights the importance of explainable and ethically responsible machine learning in educational settings.

Keywords: Machine Learning, Educational Performance, Educational Data Mining, Classification, Prediction Algorithms, Student Analytics

I. INTRODUCTION

Educational institutions generate large amounts of information through examinations, assignments, attendance records, demographic profiles, learning-management systems, online activities, and student assessments. Traditionally, these data have been analysed using descriptive statistics and conventional statistical techniques. However, the increasing volume and complexity of educational data have encouraged researchers to employ machine learning methods for discovering patterns and making predictions. Educational performance prediction generally aims to estimate outcomes such as examination scores, grade categories, course completion, academic success, or the probability that a student may experience academic difficulties.

Machine learning provides a collection of computational techniques capable of learning relationships from historical data and applying those relationships to previously unseen observations. In educational applications, predictive models may use variables such as previous academic achievement, attendance, study habits, socioeconomic characteristics, engagement, assessment scores, and online learning behaviour. The usefulness of a model depends not only on its

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, June 2026
Impact Factor: 8.2

predictive accuracy but also on interpretability, robustness, fairness, and practical usefulness for teachers and administrators.

Educational performance prediction can therefore be considered both a technical and educational problem. A highly accurate algorithm may have limited practical value if teachers cannot understand its predictions or if the model depends on biased or incomplete data. Consequently, contemporary research increasingly emphasises the balance between predictive performance and explainability.

MACHINE LEARNING APPROACHES USED FOR EDUCATIONAL PERFORMANCE PREDICTION

Machine learning algorithms applied to educational prediction can broadly be divided into supervised, unsupervised, and ensemble approaches. Supervised learning is particularly common because educational datasets frequently contain labelled outcomes such as pass/fail status, grades, or achievement levels. Classification algorithms predict categorical outcomes, whereas regression algorithms predict continuous variables such as examination marks.

Table 1. Major Machine Learning Algorithms Used for Educational Performance Prediction

Algorithm	Major Principle	Common Educational Application	Major Strength	Major Limitation
Logistic Regression	Estimates probability of categorical outcomes	Pass/fail prediction	Simple and interpretable	Limited for complex nonlinear relationships
Decision Tree	Creates rule-based splits in data	Grade or risk classification	Easy to interpret	Can overfit
Random Forest	Combines multiple decision trees	Student performance/risk prediction	Robust and generally accurate	Less interpretable than a single tree
Support Vector Machine	Finds an optimal decision boundary	Student classification	Effective in high-dimensional data	Parameter selection can be difficult
k-Nearest Neighbour	Classifies according to nearby observations	Student performance classification	Simple and intuitive	Sensitive to scaling and irrelevant variables
Naïve Bayes	Uses probabilistic classification	Academic-risk prediction	Fast and computationally efficient	Assumes conditional independence
Artificial Neural Network	Learns complex nonlinear patterns	Grade and achievement prediction	Powerful nonlinear modelling	Requires tuning and may lack transparency
Gradient Boosting	Sequentially improves weak learners	Performance and risk prediction	Strong predictive capability	Can be computationally intensive
Ensemble Methods	Combines several models	Academic outcome prediction	Often improves predictive stability	Greater complexity

Decision Trees are frequently attractive in educational applications because their decision rules can be translated into understandable conditions. Random Forest improves upon a single tree by constructing multiple trees and combining their predictions. Support Vector Machines are useful where class boundaries are complex, while k-Nearest Neighbour provides a comparatively straightforward instance-based approach.

Artificial Neural Networks are capable of modelling complicated nonlinear relationships among educational variables. However, their internal decision-making processes can be difficult to explain. Ensemble methods such as Random Forest and Gradient Boosting often provide a useful balance between predictive capability and robustness.

EDUCATIONAL VARIABLES AND DATA SOURCES FOR PREDICTION

The quality of an educational prediction model depends heavily on the quality and relevance of the variables used for training. Academic variables commonly include previous examination scores, assignment marks, internal assessment results, attendance, course grades, and progression history. Behavioural variables may include study frequency, participation in online activities, interaction with learning platforms, assignment submission patterns, and time spent on educational resources.

Demographic and socioeconomic variables can also be considered, although their use requires careful ethical evaluation. Variables such as age, educational background, location, parental education, and socioeconomic conditions may help explain differences in educational outcomes. However, predictive models should not automatically treat demographic characteristics as causal determinants of performance.

Table 2. Common Predictor Variables in Educational Performance Prediction

Variable Category	Examples	Possible Predictive Value
Academic	Previous grades, test scores, assignment marks	Very high
Attendance	Class attendance, absence frequency	High
Behavioural	Study time, participation, learning activity	High
LMS Activity	Logins, resource access, online submissions	Moderate to high
Demographic	Age, gender, educational background	Variable
Socioeconomic	Family background, economic conditions	Variable
Psychological/Learning	Motivation, engagement, learning preferences	Potentially high
Assessment	Quiz scores, continuous assessment	Very high

Data preprocessing is another critical stage. Missing observations, inconsistent records, duplicate entries, categorical variables, outliers, and imbalanced classes can affect model performance. Feature selection and feature engineering can reduce unnecessary information and improve model efficiency. Data should generally be separated into training and testing sets, with validation procedures used to reduce the possibility of overfitting.

COMPARATIVE EVALUATION OF MACHINE LEARNING ALGORITHMS

No single machine learning algorithm is universally optimal for educational performance prediction. Algorithm performance depends on the size and structure of the dataset, the number and type of predictors, the target variable, and the evaluation strategy. Accuracy is commonly reported for classification problems, but it can be misleading when classes are imbalanced. Therefore, precision, recall, F1-score, specificity, ROC-AUC, and confusion matrices can provide a more complete assessment.

For regression-based prediction, metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and coefficient of determination (R^2) can be used. Cross-validation is particularly valuable because educational datasets may be relatively small and a single train-test division may produce unstable results.

Table 3. Evaluation Measures for Educational Prediction Models

Evaluation Measure	Application	Interpretation
Accuracy	Classification	Overall proportion of correct predictions
Precision	Classification	Proportion of predicted positives that are correct
Recall	Classification	Ability to identify actual positive cases
F1-Score	Classification	Balance between precision and recall
ROC-AUC	Classification	Overall discrimination ability
MAE	Regression	Average absolute prediction error
RMSE	Regression	Penalises larger prediction errors
R^2	Regression	Proportion of explained variance

Comparative evaluation should also consider interpretability and computational requirements. Logistic Regression and Decision Trees are relatively easy to explain to educational stakeholders. Random Forest and Gradient Boosting can provide strong predictive performance, although interpretation may require feature-importance analysis or explainable-AI techniques. Neural networks can be highly effective for sufficiently large datasets but generally require greater computational resources and careful parameter tuning.

CHALLENGES, ETHICAL ISSUES AND EXPLAINABILITY

Despite their potential, machine learning applications in education face several challenges. The first is data quality. Incomplete, inaccurate, inconsistent, or unrepresentative educational records can produce unreliable predictions. A second issue is class imbalance, particularly when the objective is to identify academically at-risk students. If only a small proportion of students are at risk, a model may achieve high overall accuracy while failing to identify the students who actually need intervention.

Overfitting is another major concern. A model may perform extremely well on training data but poorly on new students. Cross-validation, regularisation, appropriate feature selection, and independent testing can help reduce this problem. Data leakage must also be avoided because information that would not realistically be available at the time of prediction can artificially inflate model performance.

Ethical considerations are equally important. Educational prediction systems may influence decisions about student support, academic intervention, or resource allocation. If historical data contain social or institutional biases, machine learning models may reproduce or amplify those biases. Student privacy and responsible data management are therefore essential. Predictions should generally be treated as decision-support information rather than unquestionable judgments about individual students.

Explainable Artificial Intelligence (XAI) can improve the usefulness of predictive models by helping educators understand why a particular prediction was generated. Feature importance, decision rules, partial dependence, and other explanation techniques can help teachers identify the variables associated with predictions. Explainability is especially important when predictions could influence interventions affecting students.

II. CONCLUSION AND FUTURE RESEARCH DIRECTIONS

Machine learning offers significant potential for educational performance prediction by enabling institutions to analyse complex patterns within academic, behavioural, attendance, demographic, and digital-learning data. Algorithms such as Decision Trees, Random Forest, Support Vector Machines, Logistic Regression, Gradient Boosting, and Artificial Neural Networks can provide useful predictive capabilities, although their effectiveness varies according to the characteristics of the dataset and prediction task.

The review indicates that algorithm selection should not be based solely on predictive accuracy. Interpretability, robustness, fairness, computational requirements, data quality, and practical educational usefulness should also be considered. Ensemble approaches may provide strong performance, while simpler models can remain valuable when transparency is a priority.

Future research should focus on explainable and fair machine learning, longitudinal student modelling, multimodal educational data, early-warning systems, personalised learning, and responsible integration of predictive models into educational decision-making. Researchers should also adopt consistent evaluation procedures and independent validation so that models developed in one educational environment can be assessed before being transferred to other institutions. Ultimately, machine learning should complement—not replace—teachers' professional judgment and should be implemented as a supportive tool for improving student learning and educational outcomes.

REFERENCES

1. ağcı, M. (2022). Educational data mining: Prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments, 9*, Article 11.
2. Asif, R., Merceron, A., Ali, S. A., & Haider, N. G. (2017). Analyzing undergraduate students' performance using educational data mining. *Computers & Education, 113*, 177–194.
3. Balcioğlu, Y. S., & Artar, M. (2025). Predicting academic performance of students with machine learning. *The Electronic Library, 41*(3).
4. Batool, S., Rashid, J., Nisar, M. W., Kim, J., Kwon, H. Y., & Hussain, A. (2023). Educational data mining to predict students' academic performance: A survey study. *Education and Information Technologies, 28*(1), 905–971.
5. Boujmiraz, S., Darhmaoui, H., & Drissi El Maliani, A. (2026). Predicting student performance: A comprehensive review of machine learning, deep learning, and explainable AI approaches. *Computers and Education: Artificial Intelligence, 10*, 100548.
6. Dabhade, P., Agarwal, R., Alameen, K. P., Fathima, A. T., Sridharan, R., & Gopakumar, G. (2021). Educational data mining for predicting students' academic performance using machine learning algorithms. *Materials Today: Proceedings, 47*(15), 5260–5267.
7. Pan, J., Zhao, Z., & Han, D. (2025). Academic performance prediction using machine learning approaches: A survey. *IEEE Transactions on Learning Technologies, 18*, 351–368.
8. Tomasevic, N., Gvozdenovic, N., & Vranes, S. (2020). An overview and comparison of supervised data mining techniques for student exam performance prediction. *Computers & Education, 143*, 103676.
9. Trakunphutthirak, R., & Lee, V. C. S. (2022). Application of educational data mining approach for student academic performance prediction using progressive temporal data. *Journal of Educational Computing Research, 60*(3).
10. 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.