
ENHANCING E-LEARNING THROUGH STUDENT PERFORMANCE PREDICTION USING ML ALGORITHMS

^{#1}Dr. PEDDI KISHOR, *Associate Professor & HOD, Department Of CSE,*

^{#2}Dr. NALLA SRINIVAS, *Associate Professor, Department Of CSE,*

^{#3}BOLLAVARAM MAHITHA, *Department Of CSE,*

SREE CHAITANYA INSTITUTE OF TECHNOLOGICAL SCIENCES, KARIMNAGAR, TG.

ABSTRACT: In the current era of data-driven decision-making, educational institutions are employing cutting-edge methodologies to assess and improve the academic performance of students. Predictive analytics, which employs machine learning methodologies, is an effective instrument for identifying patterns in educational data and forecasting future outcomes. The objective of this research is to develop a method for predicting the grades of students through the use of supervised machine learning algorithms. The dataset that was employed to develop the system encompasses academic and socioeconomic factors that influence students' performance. Data preprocessing entails the encoding of categorical variables, the removal of missing values, and the selection of critical features to improve the performance of the model. The optimal model is determined through the comparison and contrast of a variety of machine learning algorithms, including XGBoost, Linear Regression, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN). The performance of these models is evaluated using a variety of metrics and visual aids, such as MAE, MSE, R2 Score, accuracy, and confusion matrices. The model that achieved the highest R2 score was the winner. This system can be employed by educational institutions to enhance learning outcomes by predicting student performance, identifying at-risk students at an early stage, and taking the necessary action.

Keywords: *Student Performance Prediction, Machine Learning, Predictive Analytics, Supervised Learning, Educational Data Mining, Support Vector Machine (SVM), Linear Regression, XGBoost, Data Preprocessing, Feature Selection*

1. INTRODUCTION

Education is indispensable for the advancement of individuals and communities. The education sector is undergoing a significant transformation as a result of the rapid advancement of technology, and intelligence systems are being employed to improve academic outcomes and learning experiences. One of the most challenging aspects of being an academic is identifying at-risk students and providing them with the necessary assistance to enhance their performance immediately. Intricate patterns in student data are frequently overlooked by conventional evaluation techniques, despite their usefulness.

Machine learning (ML), a subset of artificial intelligence (AI), identifies concealed patterns in extensive datasets that are challenging to identify using conventional methods. It then employs these patterns to produce efficient solutions. It has the potential to transform the way we assess and improve the performance of students. Machine learning algorithms enable educational institutions to forecast students' academic performance in the future by analyzing student data, demographics, socioeconomic status, and behavioral trends. This research

introduces a supervised machine learning algorithm to predict the performance of students in class. The primary goal is to accurately predict academic performance by analyzing historical student data, employing appropriate preprocessing techniques, selecting relevant features, and training a multitude of machine learning models. The research addresses a number of critical issues, including the identification of significant influencing factors, the determination of the most effective model, and the effective interpretation of predictions by educators. These objectives are achieved by conducting a comparative analysis of a variety of supervised learning algorithms, including Linear Regression, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Extreme Gradient Boosting (XGBoost). The modular architecture of the system enables its adaptability and integration with other academic platforms or learning management systems.

This research's practical solution can be advantageous to educational institutions and academic knowledge. By utilizing predictive modeling, institutions can more effectively allocate resources, identify students at risk early, and make data-driven decisions, all of which contribute to improved academic success for students.

2. REVIEW OF LITERATURE

Namoun & Alshanqiti (2020): Conduct a comprehensive analysis of data mining and machine learning techniques for the purpose of predicting student performance in the context of online learning. The authors investigated a variety of prediction models, including neural networks, decision trees, and support vector machines. The paper's results indicate that learning analytics can be beneficial in identifying students who are likely to experience academic difficulties. It underscores the importance of educational data in improving the effectiveness of teaching strategies and the outcomes of student learning. Machine learning has the potential to significantly improve personalized education systems, according to the research.

Hwang et al. (2020): Investigates the role of AI in contemporary educational systems. The authors provide a comprehensive explanation of the ways in which AI facilitates the prediction of student performance, adaptive learning environments, and intelligent tutoring. The research reveals issues with data privacy and model interpretability, which have an impact on educational systems. It emphasizes the importance of incorporating AI into online education. The results indicate that predictive analytics have the potential to increase both student interest and academic achievement.

Albreiki et al. (2021): Provide comprehensive coverage of the use of machine learning to predict the academic performance of students. The authors compare and contrast numerous algorithms. Random forests, logistic regression, and deep learning models comprise these algorithms. The paper has identified several critical variables that have a significant impact on academic achievement, including attendance, participation, and test scores. It demonstrates the potential of machine learning to enhance educational decision-making. The results will assist e-learning platforms in delivering lessons that are more personalized for each student..

Wang et al. (2021): Propose a graph-based ensemble machine learning model. Courses, students, and learning behaviors are all examined by the system. Ensemble techniques are implemented by the authors to improve the precision and dependability of their predictions.

The model can be employed by educators to rapidly identify students who are experiencing difficulties. Intelligent data analysis is implemented to enhance the efficacy of online learning systems.

Yağcı (2022): Undertakes an examination of educational data mining methodologies to forecast students' academic performance. The author evaluates a variety of machine learning algorithms, including support vector machines, naïve Bayes, and decision trees. The investigation investigates academic and behavioral data obtained from online learning environments. Predictive models can improve the effectiveness of educational intervention and planning strategies, as demonstrated in this instance. The paper underscores the growing importance of AI-powered analytics in the field of education.

Liu et al. (2022): suggest using clickstream data from online learning systems along with machine learning to predict how well students will do in school. The writers monitor the frequency with which students access the site, view videos, and complete quizzes. The model's identified learning behaviors are associated with academic success and failure. Consequently, educators are able to offer students prompt support. The paper posits that the utilization of real-time learning analytics is essential in online learning environments.

Hussain & Khan (2023): The "Student-Performulator" is a machine learning-based framework for predicting academic performance that Hussain and Khan (2023) have proposed. Student records and behavioral data from the past are taken into account by the model. It employs classification algorithms to identify children who may experience academic difficulties. The system enables educators to implement early intervention strategies. The paper demonstrates that predictive analytics can improve student performance and educational quality.

Ahmed et al. (2024): Implement numerous machine learning algorithms to forecast students' academic performance. The authors assess a variety of models, such as neural networks, decision trees, and random forests, in order to identify the most effective one. The system evaluates the results of exams, homework, and class attendance. This tool enables educators to more effectively monitor the academic development of their students. The research's primary focus is on the utilization of AI to improve student achievement and online learning.

Al-Tameemi et al. (2024): A hybrid ML approach is recommended by Al-Tameemi et al. (2024) to forecast students' academic performance. The accuracy of predictions is improved while error rates are reduced by the integration of multiple algorithms in the model. It processes extensive educational datasets that are obtained from online learning platforms. The system also supports personalized learning strategies and identifies students who are not performing well. The paper demonstrates that hybrid AI models can be advantageous for educational analytics.

Fazil et al. (2024): In order to anticipate the success of students in virtual classrooms, Fazil et al. (2024) propose a deep learning algorithm. Neural networks are implemented by the authors to evaluate behavioral and academic data. The system is accurate in its ability to forecast academic success in the future. Consequently, educators are able to customize their lessons to meet the unique requirements of their students. The paper underscores the importance of deep learning in intelligent e-learning environments.

Arévalo-Cordovilla & Peña (2024): Compare machine learning models for predicting student performance in online programming courses. The authors evaluate the precision and efficacy of various algorithms. The primary objective of the research is to determine the factors that influence the performance of students in online courses. It can be employed by educators to improve the development of curriculum and class participation. The findings enable the implementation of flexible online education.

Wang (2025): This research presents a machine learning approach to predicting the performance of students in online classes (Wang, 2025). The model is fed with data regarding student interactions, test scores, and participation rates. It identifies patterns that may indicate academic success or failure. The system has the capacity to issue early warnings to students who may require academic assistance. This research demonstrates that predictive analytics can enhance the efficacy of online courses.

Ahmed et al. (2025): Initiated a project to forecast students' academic performance using machine learning in order to provide outstanding education. Complex AI algorithms are implemented by the authors to analyze educational datasets. This system enables institutions of higher education to monitor student development and identify areas of weakness. It enables educators to make data-driven decisions and plan. The significance of AI technology in the development of personalized and effective learning environments is underscored by the research.

Boujmiraz (2026): The extensive research conducted by Boujmiraz (2026) is focused on the use of deep learning, machine learning, and explainable AI to forecast student performance. The author explores the subject of educational analytics and the influence of contemporary AI methodologies. The primary focus of the investigation is the issues of model interpretability, fairness, and transparency. It emphasizes the importance of AI that is comprehensible in order to foster confidence in educational institutions. The research promotes the use of e-learning platforms that are both transparent and intelligent.

Sonkhanani et al. (2026): Develop a machine learning model for predicting student grades by utilizing demographic data and academic records. The system considers a variety of factors that influence students' academic performance. Predictive analytics is employed to identify students who are likely to experience academic difficulties. The model enables educators to implement personalized intervention strategies. The research indicates that educational management systems that are powered by AI have significant potential.

3. PROPOSED METHODOLOGY

The proposed work endeavors to develop an intelligent system that can anticipate the academic performance of students in the future by utilizing historical academic and demographic data. The system is not only practical in educational contexts, but it is also scalable and user-friendly. Additionally, it generates accurate forecasts. Detailed in this section are the system's behavior under various operating conditions, design considerations, and functional objectives.

The system endeavors to predict academic outcomes, such as grades, for students by utilizing a variety of factors. Early identification of at-risk students, customization of learning

strategies, and more efficient allocation of resources are all facilitated by the system's real-time insights into student performance.

- User Interface for File Upload:**
 Students can effortlessly upload CSV datasets through an intuitive, Streamlit-powered user interface.
- Data Format Validation:**
 The data is automatically cleaned, encoded, and normalized using the same preprocessing methods that were employed to train the model..
- Preprocessing Pipeline:**
 The data is automatically cleaned, encoded, and normalized using the same preprocessing methods that were employed to train the model.
- Model Selection Logic:**
 Models including SVM, KNN, Linear Regression, and XGBoost are assessed, and the model with the highest R^2 score is chosen.
- Prediction Output:**
 A new column named "Prediction" is incorporated into the dataset to forecast students' grades, and the selected model is employed.
- Downloadable Results:**
 We offer a final output file in CSV format that includes the predictions for your analysis and reporting requirements.

4. RESULTS

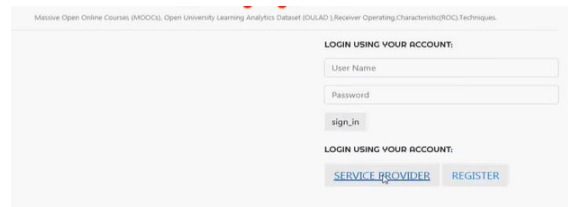


Fig 1: Service Provider Login

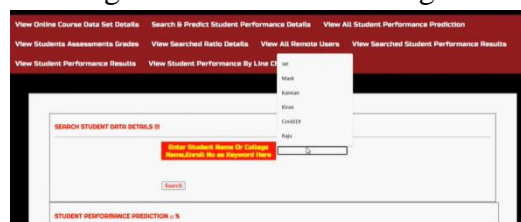
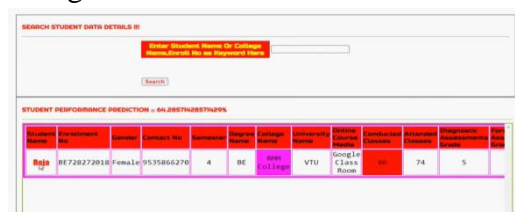


Fig 2: Search Student Data Details



Student ID	Roll Number	Name	College Name	Branch	Year	Grade	Score	Prediction
101	RE728272010	Female	SVS3866270	4	BE	90	90	74

Fig 3 : Student Performance Prediction

4. Wang, Y., Ding, A., Guan, K., Wu, S., & Du, Y. (2021). Graph-based Ensemble Machine Learning for Student Performance Prediction. arXiv preprint arXiv:2112.07893.
5. Yağcı, M. (2022). Educational Data Mining: Prediction of Students' Academic Performance Using Machine Learning Algorithms. *Smart Learning Environments*, 9(11), 1–19.
6. Liu, Y., Fan, S., Xu, S., Sajjanhar, A., Yeom, S., & Wei, Y. (2022). Predicting Student Performance Using Clickstream Data and Machine Learning. *Education Sciences*, 13(1), 17.
7. Hussain, S., & Khan, M. Q. (2023). Student-Performulator: Predicting Students' Academic Performance at Secondary and Intermediate Level Using Machine Learning. *Annals of Data Science*, 10(3), 637–655.
8. Ahmed, E., et al. (2024). Student Performance Prediction Using Machine Learning Algorithms. *Hindawi Scientific Programming*, 2024, 4067721.
9. Al-Tameemi, G., et al. (2024). A Hybrid Machine Learning Approach for Predicting Student Academic Performance. *Procedia Computer Science*, 245, 344–352.
10. Fazil, M., et al. (2024). A Novel Deep Learning Model for Student Performance Prediction. *Journal of Learning Analytics*, 11(2), 44–58.
11. Arévalo-Cordovilla, F. E., & Peña, M. (2024). Comparative Analysis of Machine Learning Models for Predicting Student Success in Online Programming Courses. *Mathematics*, 12(20), 3272.
12. Wang, J. (2025). Machine Learning Approach to Student Performance Prediction Based on Online Learning. *Scientific Reports*, 15, 11731722.
13. Ahmed, W., et al. (2025). Machine Learning-Based Academic Performance Prediction for Quality Education Delivery. *Scientific Reports*, 15, 12353.
14. Boujmiraz, S. (2026). Predicting Student Performance: A Comprehensive Review of Machine Learning, Deep Learning, and Explainable AI Approaches. *Computers and Education: Artificial Intelligence*, 10, 100548.
15. Sonkhanani, M., Chibaya, S., & Nyirenda, C. N. (2026). Machine Learning Grade Prediction Using Students' Grades and Demographics. arXiv preprint arXiv:2603.00608.