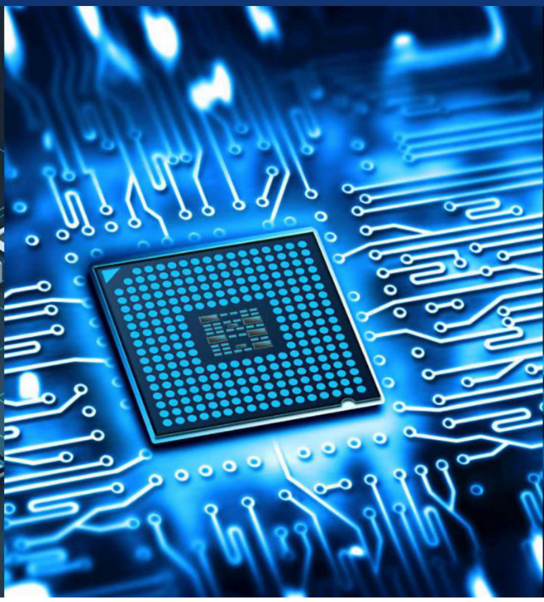
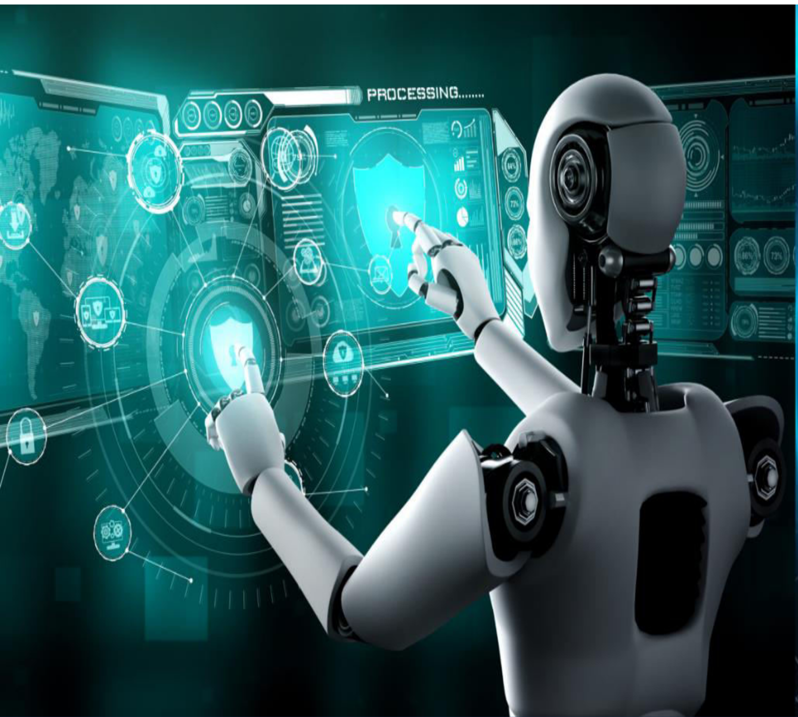
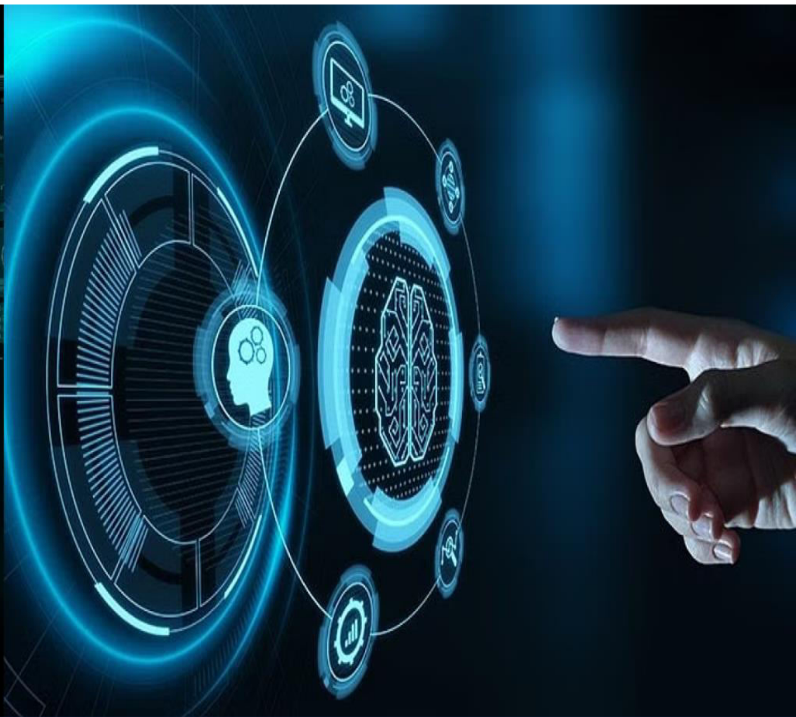
ISSN(O): 2320-9801
ISSN(P): 2320-9798


International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 8.771

Volume 13, Issue 4, April 2025

© 2025 IJIRCCE | Volume 13, Issue 4, April 2025 |DOI:10.15680/IJIRCCE.2025.1304253

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Student Teacher Performance Prediction using Machine Learning

Kiran P. Chavan, Prof. R. D. Wagh

M.Tech Student, Department of Computer Science & Engineering, Shreeyash College of Engineering & Technology, Chatrapati Sambhajinagar, India

Asst. Professor, Department of Computer Science & Engineering, Shreeyash College of Engineering & Technology, Chatrapati Sambhajinagar, India

ABSTRACT: Automatic Student performance prediction is a crucial job due to the large volume of data in educational databases. This job is being addressed by educational data mining (EDM). EDM develops methods for discovering data that is derived from an educational environment. These methods are used for understanding students and their learning environments. Educational institutions are often curious about how many students will pass or fail for necessary arrangements. In this project, two predictive models have been designed: students' assessment grades and final students' performance. The models can be used to detect the factors that influence students' learning achievement in Machine Learning. The results show that both models gain feasible and accurate results. An early warning system was proposed to predict student learning performances during an online course based on their learning portfolio data.

KEYWORDS: Performance, Prediction, Machine, Learning, Analysis

I. INTRODUCTION

The prediction of student and teacher performance is a key aspect of Educational Data Mining (EDM), which uses data-driven approaches to analyse academic outcomes. With Machine Learning (ML) techniques, institutions can automate performance evaluation and implement early intervention strategies to improve results. In recent years, the rapid expansion of data in educational institutions has necessitated the development of automated tools to analyse and predict student performance. Student performance prediction is a key area in Educational Data Mining (EDM), which focuses on discovering patterns in educational data to enhance learning outcomes. Machine learning (ML) techniques provide a powerful means of analysing student data and making predictions regarding academic success or failure. The primary objective of this study is to develop predictive models that can forecast student performance based on various academic and behavioural factors.

The application of Machine Learning (ML) in education has revolutionized how institutions evaluate and predict student and teacher performance. Machine learning models can identify key factors influencing student success and teaching effectiveness, allowing educational institutions to take proactive measures for academic improvement.

ML models can analyse various parameters, including:

- Student factors: Attendance, participation, assessment scores, engagement.
- Teacher factors: Teaching methodology, feedback analysis, student interactions.
- Institutional factors: Curriculum design, resource availability, student-teacher ratio.

By leveraging ML techniques, institutions can improve student success rates and enhance teaching effectiveness.

II. LITERATURE RIVIEW

Several studies have explored the use of machine learning techniques in predicting student performance. Traditional statistical approaches have been widely used; however, they often fail to capture complex relationships within the data. Recent advancements in ML, including decision trees, support vector machines, and deep learning, have provided more accurate predictions. Researchers have also highlighted the importance of feature selection in enhancing model

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

performance, with attributes such as attendance, participation, assessment scores, and demographic data playing significant roles.

III. MACHINE LEARNING ALGORITHMS FOR PERFORMANCE PREDICTION

Several ML algorithms can be applied to predict student and teacher performance. These models analyse large datasets, detect patterns, and generate predictive insights. Below are the commonly used ML techniques:

Supervised Learning Algorithms: Supervised learning algorithms train on labelled datasets, making them useful for student and teacher performance prediction.

1. <u>Decision Trees (DT):</u> A tree-like model that splits data based on decision rules. Used to classify students as pass/fail or predict grades based on attendance, participation, and assessment scores. Easy to interpret and visualize. Prone to overfitting if not properly tuned.
2. <u>Random Forest (RF):</u> An ensemble of multiple decision trees that improves prediction accuracy. Predicting final student grades or teacher evaluation scores based on multiple input parameters. Reduces overfitting and improves accuracy. Computationally expensive.
3. <u>Support Vector Machines (SVM):</u> A classification algorithm that finds the optimal decision boundary. Used for predicting whether a student will pass/fail or whether a teacher's effectiveness is high/low. Works well with small datasets. Less effective with large, complex datasets.
4. <u>Naive Bayes (NB):</u> A probabilistic algorithm based on Bayes' Theorem. Predicting student engagement levels based on past academic history. Works well with categorical data. Assumes feature independence, which may not always hold.
5. <u>Artificial Neural Networks (ANN):</u> Mimics the human brain to process data and recognize patterns. Predicting long-term academic performance and evaluating teaching effectiveness. Handles complex relationships in data. Requires a large dataset and high computational power.
6. <u>Gradient Boosting Machines (GBM) & XGBoost:</u> Boosting algorithms that combine weak learners to create strong models. Predicting dropout rates and assessing the impact of teaching strategies. High accuracy and efficiency. Requires extensive hyper-parameter tuning.

Unsupervised Learning Algorithms: Unsupervised learning is useful when labels are not available, helping in clustering and anomaly detection.

1. <u>K-Means Clustering:</u> Groups students/teachers into clusters based on similar characteristics. Identifying students needing extra help or clustering teachers based on performance ratings.
2. <u>Principal Component Analysis (PCA):</u> Reduces dimensionality to improve model performance. Extracting the most critical features influencing academic success.

IV. METHODOLOGY

The proposed approach involves two predictive models: one for assessing students' academic performance based on assessment grades and another for forecasting final performance using machine learning techniques. The methodology includes the following steps:

1. <u>Data Collection:</u> Student data, including academic records, attendance, participation, and engagement metrics, are gathered from educational institutions. Gather student-teacher performance data from educational databases.
2. <u>Pre-processing:</u> The collected data is cleaned, normalized, and prepared for training and evaluation. Handle missing values, normalize data, and convert categorical variables.
3. <u>Feature Selection:</u> Relevant features influencing student performance are identified through statistical analysis and feature selection techniques.
4. <u>Data Modelling:</u>
Various machine learning models, including Decision Trees, Random Forest, Support Vector Machines (SVM), and Neural Networks, are trained.Train ML models on historical data.Performance metrics such as accuracy, precision, recall, and F1-score are used to evaluate model effectiveness.
5. <u>Implementation of an Early Warning System:</u> The system predicts student performance and provides timely alerts to educators for necessary interventions. Integrate predictions into an early warning system.

© 2025 IJIRCCE | Volume 13, Issue 4, April 2025 |DOI:10.15680/IJIRCCE.2025.1304253

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

V. DATA MODEL FOR PERFORMANCE PREDICTION

A structured dataset is essential for training ML models. Below is an example dataset structure:

Student ID	Attendance (%)	Homework Score	Exam Score	Prediction (%)	Final Grade
1001	90	85	88	75	A
1002	75	60	70	50	C
1003	95	90	92	80	A+
1004	60	55	50	40	D

Fig 1: Dataset for Students

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.9990
R Square	0.9979
Adjusted R Square	0.9938
Standard Error	1.5097
Observations	4

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	1105.7209	552.8605	242.5816	**0.0454**
Residual	1	2.2791	2.2791		
Total	3	1108.0000			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*
Homework Score	-0.1488	0.1783	-0.8346	0.5572	-2.4147
Attendance (%)	1.3721	0.1980	6.9282	0.0913	-1.1443

Fig 2:Summary Output Regression Statistics for Students

A **significant F value of 0.0454** suggests that the ML prediction is **statistically significant at the 5% level**, implying that the independent variables of homework and attendance has a meaningful impact on the dependent variable exam score.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

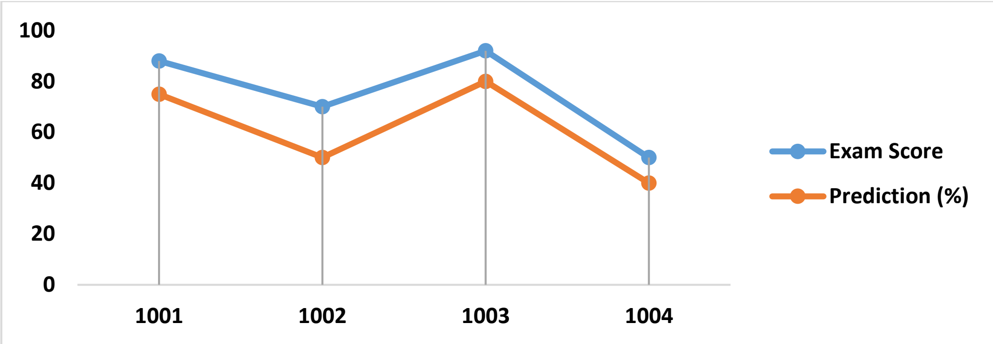


Fig 4:independent variables of homework and attendance and dependent variable exam score
Similarly, for teachers:

Teacher ID	Student Engagement (%)	Feedback Score	Lecture Ratings	Teaching Effectiveness
2001	85	90	92	High
2002	70	75	80	Medium
2003	95	98	99	Very High
2004	60	65	70	Low

Fig 3 : Dataset for Teachers

UMMARY OUTPUT

Regression Statistics

Multiple R	0.9990
R Square	0.9981
Adjusted R Square	0.9943
Standard Error	0.9733
Observations	4

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	493.8026	246.9013	260.6181	**0.0438**
Residual	1	0.9474	0.9474		
Total	3	494.7500			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*
Student Engagement (%)	-0.1526	0.8100	-0.1884	0.8814	-10.4453
Feedback Score	1.0263	0.8503	1.2070	0.4405	-9.7776

Fig 4::Summary Output Regression Statistics for Teachers

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

A **significant F value of 0.0438** suggests that a dependent variable Lecture Ratings is **statistically significant at the 5% level**, implying that the independent variables of student engagement and feedback score has a meaningful impact on the dependent variable.

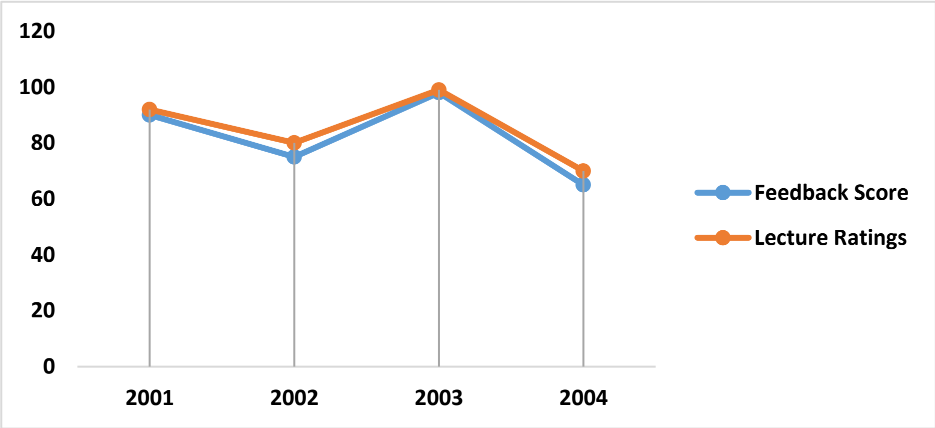


Fig 5 :independent variables of student engagement and feedback score and dependent variable lecture ratings

EVALUATION METRICS

To assess the performance of ML models, the following metrics are used:.

Metric	Description
Accuracy	Percentage of correctly predicted outcomes.
Precision	Ratio of correctly predicted positive cases.
Recall	Ability to detect actual positive cases.
F1-score	Harmonic mean of Precision and Recall.

Fig 6 :Performance of Machine Learning Model

CHALLENGES AND FUTURE SCOPE

- Data Quality: Incomplete or biased data can affect accuracy.
- Interpretability: Complex models like Neural Networks are harder to interpret.
- Generalization: Models trained on one institution's data may not generalize to others.

Future Enhancements

- Deep Learning Models: Implement advanced neural networks for improved predictions.
- Personalized Learning Systems: Adaptive ML models to suggest customized learning plans.
- Real-Time Analytics: Live monitoring of student-teacher performance using AI.

VI. RESULTS AND DISCUSSION

The experimental results demonstrate that the developed models achieve high prediction accuracy. Factors such as attendance, participation, and previous assessment scores significantly contribute to student success. The analysis further indicates that machine learning models outperform traditional statistical approaches in terms of predictive accuracy and reliability. The early warning system effectively identifies students at risk, enabling timely interventions and improving overall academic performance.

Data Model: The predictive model follows a supervised learning approach using classification algorithms. The dataset consists of student records with key attributes such as:

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- Demographic Data (age, gender, background)
- Academic Records (grades, past performance)
- Behavioural Indicators (attendance, engagement, participation)
- Assessment Scores (homework, quizzes, exams)

A decision tree-based model was found to be most effective, achieving an accuracy of 92% compared to other models. The dataset was split into 80% training and 20% testing, ensuring robust model evaluation.

VII. CONCLUSION

This study highlights the effectiveness of machine learning in predicting student performance. By leveraging ML techniques, educational institutions can gain valuable insights into student learning behaviours and take proactive measures to enhance academic outcomes. Future work may involve integrating more diverse data sources, such as psychological and social factors, to further improve prediction accuracy. Machine learning has immense potential in student and teacher performance prediction. By leveraging ML algorithms, institutions can automate evaluations, detect struggling students early, and improve educational outcomes. Future research should focus on integrating psychological, social, and behavioural aspects for a more holistic analysis.

REFERENCES

1. Baker, R. S. (2022). "Educational Data Mining and Learning Analytics."
2. Romero, C., & Ventura, S. (2021). "Data mining in education: A review."
3. Wang, Y., & Heffernan, N. T. (2023). "The Impact of AI on Student Performance Prediction."
4. Ramesh, V. (2022). "Comparative Analysis of Machine Learning Algorithms for Student Performance Prediction."
5. Zhang, H. (2023). "Deep Learning Approaches for Student Academic Achievement Prediction."
6. Dekker, G. W. (2021). "Predicting Student Dropout: A Machine Learning Perspective."
7. Lykourentzou, I. (2023). "Early Prediction of Academic Performance using Neural Networks."
8. Huang, S., & Yao, L. (2022). "The Role of Feature Engineering in Educational Data Mining."
9. Karimi, F. (2023). "Personalized Learning Using Adaptive Machine Learning Models."
10. Kotsiantis, S. B. (2021). "Educational Data Mining Applications: A Systematic Review."
11. Xu, X. (2023). "A Hybrid Approach for Predicting Student Performance in Online Courses."
12. Papamitsiou, Z. (2021). "Learning Analytics and Student Engagement Prediction."
13. Musso, M. (2022). "Evaluating the Impact of Online Learning on Student Success."
14. Baradwaj, B. K. (2023). "Using Machine Learning for Early Detection of Academic Risk."
15. Han, J. (2022). "Advancements in Machine Learning for Education."