

Students Performance Prediction System Using Machine Learning

Y Shaheela¹, S. Vigashini², Av. Surya Prakash³, K. Thameema⁴, Mohammed A⁵, Suresh kumar A⁶

¹Assistant Professor, Department of Information Technology, Rathinam Technical Campus, Eachanari, Coimbatore, Tamil Nadu, India.

^{2,3,4,5}Department of Information Technology, Rathinam Technical Campus, Eachanari, Coimbatore, Tamil Nadu, India.

⁶Assistant Professor, Department of CSE, Rathinam Technical Campus Coimbatore, Tamil Nadu, India.

OPEN ACCESS

Article Citation:

Y Shaheela¹, S. Vigashini², Av. Surya Prakash³, K. Thameema⁴, Mohammed A⁵, Suresh kumar A⁶, "Students Performance Prediction System Using Machine Learning", International Journal of Recent Trends in Multidisciplinary Research, July-August 2026, Vol 6(04), 438-441.



©2026 The Author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. Published by 5th Dimension Research Publication

Abstract: Student performance prediction has become an important research area in educational data mining. This project proposes a Machine Learning-based system that analyzes academic and behavioral data to predict student performance in advance. The system uses algorithms such as Decision Tree, Random Forest, and Logistic Regression to classify students into performance categories (High, Medium, and Low) or predict final grades. The dataset includes attributes such as attendance, internal marks, assignment scores, study hours, and previous semester results. The trained model identifies at-risk students early and provides actionable insights for teachers and institutions. The proposed system improves academic monitoring, enables early intervention, and enhances overall student success rates through data-driven decision making.

Keywords: Machine Learning, Predictive Analytics, Academic Performance, Data Mining Classification Algorithms.

1. Introduction

In recent years, Machine Learning and Data Analytics have significantly transformed various sectors including healthcare, finance, agriculture, and education. In the education sector, predicting student performance has gained increasing importance to improve learning outcomes and reduce failure rates.

Traditional evaluation systems assess student performance only after examinations.

However, early prediction of academic outcomes can help teachers take preventive measures before final exams.

The proposed Student Performance Prediction System uses machine learning techniques to analyze student academic data and predict their performance level.

2. Literature Survey

1. Cortez and Silva (2008) developed a student performance prediction model using Decision Trees and Neural Networks. Their work demonstrated that student grades can be predicted using demographic and academic attributes.
2. Kotsiantis et al. (2010) applied classification algorithms such as Naïve Bayes and Decision Trees to predict academic performance. The study highlighted the importance of attendance and internal assessments.
3. Pandey and Pal (2011) used data mining techniques to identify factors influencing student performance and concluded that study habits significantly affect academic success.
4. Huang and Fang (2013) utilized Logistic Regression and Support Vector Machines (SVM) to predict student outcomes and found that early academic indicators strongly influence final grades.
5. Ahmad et al. (2015) implemented Random Forest and Neural Networks for performance prediction and reported improved accuracy compared to traditional statistical methods.

3. Problem Statement

In the modern educational environment, academic success is influenced by multiple interconnected factors such as attendance, internal assessments, assignment performance, study habits, participation level, socio-economic background, and prior academic records. However, educational institutions often lack an intelligent mechanism to analyze these factors

Students Performance Prediction System Using Machine Learning

collectively and predict student outcomes in advance.

Student underperformance is typically identified only after final examinations, by which time corrective action becomes difficult or ineffective. Early identification of academically at-risk students remains a major challenge for schools and colleges. Without predictive analysis, educators rely mainly on manual observation and intuition, which may not accurately reflect the student's overall academic trajectory.

Educational institutions face challenges in identifying academically weak students before final examinations.

Current systems:

- Depend on manual analysis
- Lack predictive capability
- Do not provide early warning alerts

Student performance depends on multiple factors

- Study hours
- Previous academic record.
- Attendance percentage.
- Internal Assignment marks
- Assignment Completion

4. Methodology

The proposed methodology follows a structured Machine Learning workflow.



Figure: Proposed System Architecture

Figure 1: High-Level Project Workflow

A. Data Collection

The dataset is collected from institutional academic records.

Attributes include:

- Attendance percentage
- Internal marks
- Assignment scores
- Study hours
- Previous semester GPA
- Participation level

B. Data Preprocessing

- Handling missing values
- Data normalization
- Feature selection
- Encoding categorical variables.

C. Feature Selection

Important features influencing performance are selected using correlation analysis and feature importance techniques.

D. Model Training

Machine Learning algorithms used:

- Decision Tree
- Random Forest

5. Implementation

A. System Architecture

The system consists of:

1. Machine Learning Model
2. Prediction Module
3. Dashboard Interface
4. Preprocessing Module



B. Model Design

The classification model follows this structure:

Input Layer → Feature Processing → ML Algorithm
→ Output Prediction

C. Training and Optimization

The model uses optimization techniques such as:

- Gradient Descent
- Hyper parameter tuning
- Cross-validation.

D. User Interface

The web interface allows:

- Admin to upload dataset
- Teacher to view predictions
- Display of risk-level students
- Visualization through graphs

6. Experimental Results

The trained model was tested using unseen student data.

Model Accuracy Graph

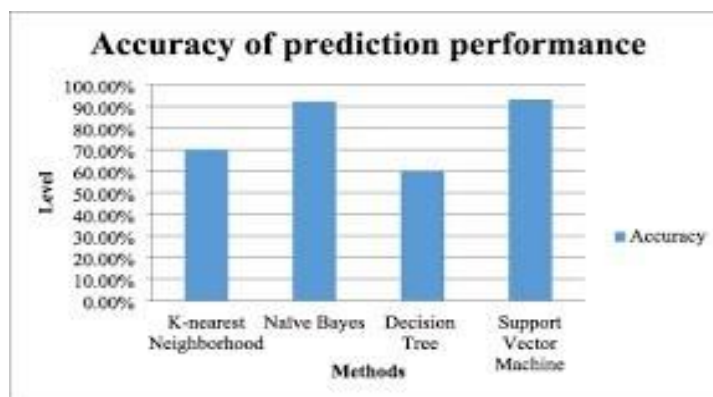
Training accuracy increased steadily with epochs, indicating proper learning without significant overfitting.

Model Loss Graph

Loss decreased gradually, confirming optimization efficiency.

Confusion Matrix

The confusion matrix shows correct classification of High, Medium, and Low performing students with minimal misclassification.



7. Conclusion

The Student Performance Prediction System presented in this work demonstrates the effective application of Machine Learning techniques in the field of education. By analyzing academic and behavioral attributes such as attendance, internal marks, assignment performance, study hours, and previous academic records, the system successfully predicts student performance levels in advance.

The developed model provides a structured, data-driven approach to academic evaluation rather than relying solely on traditional examination-based assessments. The experimental results show that classification algorithms such as Random Forest and Decision Tree achieve high prediction.

One of the major contributions of this work is the early identification of academically weak students, enabling timely intervention strategies such as counseling, remedial classes, and personalized mentoring. This proactive approach can significantly reduce failure rates, improve student confidence, and enhance overall academic performance.

Additionally, the system promotes data-driven decision-making within educational institutions. Administrators and faculty members can use predictive insights to design better academic policies, optimize teaching methodologies, and allocate resources efficiently.

Overall, the proposed system bridges the gap between educational data and intelligent analytics, making academic monitoring more efficient, accurate, and scalable.

References

1. P. Cortez and A. Silva, "Using Data Mining to Predict Secondary School Student Performance," in Proceedings of 5th Future Business Technology Conference, 2008.
2. S. B. Kotsiantis, C. J. Pierris, and P. E. Pintelas, "Predicting Students' Performance in Distance Learning Using Machine Learning Techniques," *Applied Artificial Intelligence*, vol. 18, no. 5, pp. 411–426, 2010.
3. U. K. Pandey and S. Pal, "Data Mining: A Prediction of Performer or Underperformer Using Classification," *International Journal of Computer Science and Information Security*, vol. 9, no. 4, pp. 136–140, 2011.
4. S. Huang and N. Fang, "Predicting Student Academic Performance in an Engineering Dynamics Course: A Comparison of Four Types of Predictive Mathematical Models," *Computers & Education*, vol. 61, pp. 133–145, 2013.
5. F. Ahmad, N. H. Ismail, and A. A. Aziz, "The Prediction of Students' Academic Performance Using Classification Data Mining Techniques," *Applied Mathematical Sciences*, vol. 9, no. 129, pp. 6415–6426, 2015.
6. C. Romero and S. Ventura, "Educational Data Mining: A Review of the State of the Art," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 40, no. 6, pp. 601–618, 2010.
7. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed., Springer, 2009.
8. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
9. D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *arXiv preprint arXiv: 1412.6980*, 2014.
10. S. Kotsiantis, "Use of Machine Learning Techniques for Educational Proposes: A Decision Support System for Forecasting Students' Grades," *Artificial Intelligence Review*, vol. 37, no. 4, pp. 331–344, 2012.