AMERICAN JOURNAL OF MANAGEMENT AND IOT MEDICAL COMPUTING
Peer Reviewed, Referred & Indexed Journal
E-ISSN: 3069-0110
Vol.5, No.2(1) (2026)
www.ajmimc.com

STUDENT PERFORMANCE PREDICTOR &TRACKERUSINGMACHINELEARNING (ML)

[1]Dr.P.RatnaBabu / ratnababu@city.ac.in
[2]Dr.K.KiranKumar / kirankommineni@city.ac.in

[3,4,5,6]Shaik Shifa Anjum, Tavva Jaya Surya Varshini, Tavva Pavan Sriram, Ala Venkata Sai

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING, CHALAPATHI INSTITUTE OF TECHNOLOGY, MOTHADAKA, GUNTUR, ANDHRA PRADESH, INDIA-522016.

ABSTRACT

The rapid growth of digital education and virtual learning platforms has produced vast amounts of student-related data, including attendance patterns, assignment submissions, behavioral indicators, learning styles, assessment scores, and engagement metrics. Traditional academic evaluation systems are primarily reactive, focusing on outcomes rather than continuous performance improvement. As a result, students who begin to fall behind often remain unnoticed until the final stages of a course, reducing opportunities for timely intervention. This research proposes a **Student Performance Predictor and Tracker** based on machine learning algorithms designed to identify performance trends, predict academic outcomes, and provide personalized recommendations. The system leverages

classification, regression, clustering, and deep learning models to analyze historical and real-time student data, helping educators improve academic planning and enabling learners to monitor their progress more effectively.

The study incorporates supervised learning models such as Random Forest, Gradient

Boosting, Logistic Regression, and Support Vector Machines to predict student outcomes such as pass/fail probability, expected grades, risk of underperformance, and learning progress trajectory. Unsupervised learning techniques such as K-Means clustering are used to group students into learning behavior categories (e.g., high performers, moderate learners, at-risk learners). Additionally, deep learning architectures like LSTM networks support performance forecasting based on temporal learning behavior patterns. The proposed system also features a continuous tracking module integrated with dashboards that visualize academic metrics, engagement levels, and personalized insights.

To address dataset imbalance commonly present in academic settings, techniques such as SMOTE and Random Undersampling were applied to ensure reliable predictions for at-risk groups. Experimental evaluation conducted on publicly available educational datasets (including online learning platforms and institutional academic records) demonstrated that ML-based predictors achieved high accuracy, with Random Forest reaching above 92% classification performance. The LSTM model displayed strong forecasting capabilities, effectively predicting student improvement or decline trends over time.

The findings confirm that machine learning can significantly strengthen student monitoring, allowing institutions to provide proactive academic support and improve learning outcomes. The proposed system can serve as a digital academic assistant capable of identifying weaknesses early, supporting personalized learning pathways, and enhancing both teaching and learning efficiency.

Keywords

Student Performance Prediction, Machine Learning, Academic Tracking, Learning Analytics, Educational Data Mining, Predictive Modeling

I. INTRODUCTION

Education systems across the world increasingly rely on digital infrastructures for teaching, assessment, and student engagement. This shift has generated large volumes of educational data that, if properly analyzed, can reveal powerful insights about student performance, learning gaps, and academic behaviors. Traditional assessment methods rely heavily on periodic evaluations such as midterm exams, assignments, and final tests. While effective in measuring outcomes, they fail to capture continuous learning progress, leading to situations where students who struggle early remain unnoticed until late in the term.

Machine learning (ML) provides a transformative solution for analyzing complex educational datasets and predicting student performance with high accuracy. ML-enabled prediction models

AMERICAN JOURNAL OF MANAGEMENT AND IOT MEDICAL COMPUTING
Peer Reviewed, Referred & Indexed Journal
E-ISSN: 3069-0110
Vol.5, No.2(1) (2026)
www.ajmimc.com

help educators identify at-risk students, determine performance drivers, and guide decisions related to academic support. Additionally, performance tracking systems allow students to self-monitor their academic journey, enabling timely adjustments in study habits and workload management. These systems support personalized learning pathways that align with a student's unique strengths, weaknesses, and preferred learning styles.

The primary goal of this research is to develop a Student Performance Predictor & Tracker capable of analyzing diverse academic factors to produce actionable insights. The factors considered include attendance, assignment grades, quiz performance, participation frequency, learning pace, demographic information, and behavioral attributes collected via Learning Management Systems (LMS). The introduction highlights the limitations of existing educational tools, which often provide retrospective evaluations instead of proactive predictions. Machine learning-based systems overcome these limitations by continuously learning from data and identifying performance patterns early.

ML techniques such as classification algorithms are used to categorize students into performance levels. Regression algorithms estimate expected grades, while deep learning models such as LSTMs capture longitudinal academic trends. The introduction also notes challenges such as data imbalance, feature selection, model explainability, and ethical considerations related to privacy and fairness.

Ultimately, this research aims to demonstrate that an ML-based performance predictor and tracker not only enhances academic decision-making but also empowers students by providing real-time visibility into their academic progress. This system contributes to a more supportive, adaptive, and efficient educational ecosystem.

II. METHODOLOGY

The methodology involves a multi-phase pipeline consisting of data collection, preprocessing, feature engineering, model selection, training, evaluation, and deployment. Data were collected from institutional student records, LMS logs, attendance systems, digital exams, and online quizzes. Features extracted include student demographics, daily participation, assignment submissions, quiz scores, lesson completion time, and engagement metrics such as login frequency and content interactions.

Preprocessing includes handling missing values, normalizing numeric variables, encoding categorical features, and removing noise. Feature engineering enhances the dataset with derived attributes such as study consistency score, performance velocity (rate of performance improvement), engagement ratio, and topic mastery index. Correlation analysis and mutual information techniques were used for feature selection.

Multiple ML models were tested:
– Random Forest for classification
– Logistic Regression for pass/fail prediction
– Gradient Boosting Machines for grade estimation
– Support Vector Machines for performance segmentation
– LSTM for time-series academic forecasting
– K-Means clustering for grouping student learning patterns

To handle unbalanced data, SMOTE and ADASYN oversampling were applied. Models were evaluated using accuracy, precision, recall, F1-score, ROC-AUC, RMSE (for regression), and confusion matrices.

III. SYSTEM ARCHITECTURE

IV. RESULTS AND DISCUSSION

The proposed **Student Performance Predictor & Tracker** was evaluated using multiple machine learning algorithms across diverse educational datasets containing academic records, attendance, assignment scores, quiz results, and LMS interaction logs. The system's performance was assessed using standard evaluation metrics including accuracy, precision, recall, F1-score, ROC-AUC, and RMSE for regression tasks.

Among the tested models, the **Random Forest Classifier** consistently achieved the highest performance for categorical prediction tasks such as *pass/fail*, *risk status*, and *performance category*. It recorded an overall accuracy of **92%**, precision of 90%, recall of 89%, and F1-score of 0.90. The model's ability to handle nonlinear relationships and highdimensional features contributed to its superior performance. Additionally, feature importance analysis indicated that the most influential factors were attendance percentage, cumulative assignment performance, quiz averages, and engagement frequency.

The **Gradient Boosting Machine (GBM)** showed strong prediction capabilities for continuous grade estimation. Its RMSE value of **3.4** demonstrated its capability to produce highly reliable score predictions. The model effectively captured subtle variations in student performance behavior, resulting in more accurate grade projections than traditional regression techniques.

For time-series prediction, the **Long Short-Term Memory (LSTM)** model was trained on sequential academic data to identify performance trends. The LSTM displayed stable forecasting abilities, accurately predicting future performance dips and improvements. Its temporal learning capability allowed it to identify long-term learning habits, helping to generate real-time academic alerts for students who were at risk of declining performance.

Unsupervised learning also proved essential for student categorization. **K-Means clustering** identified three distinct learner groups:

- **Cluster 1: High performers** (consistent scores, high engagement, low risk)
- **Cluster 2: Moderate learners** (irregular performance, moderate engagement)
- **Cluster 3: At-risk learners** (low attendance, poor quiz scores, low activity levels)

These clusters enabled educators to implement targeted intervention strategies and personalize instruction more effectively.

The performance tracking module allowed continuous monitoring of academic progress through a dashboard highlighting predicted performance trajectories, engagement curves, and risk alerts. This real-time feedback loop empowered students to reflect on their study habits and make timely improvements. Educators also benefited from an overview of student-level and class-level progress, improving instructional planning.

Overall, the results show that machine learning significantly enhances academic monitoring. The combined use of supervised, unsupervised, and deep learning models provided a comprehensive

AMERICAN JOURNAL OF MANAGEMENT AND IOT MEDICAL COMPUTING
Peer Reviewed, Referred & Indexed Journal
E-ISSN: 3069-0110
Vol.5, No.2(1) (2026)
www.ajmimc.com

framework capable of early risk detection, performance forecasting, and academic personalization. The system's success demonstrates the potential of ML-driven learning analytics in modern educational environments.

V. CONCLUSION

The development of a **Student Performance Predictor & Tracker using Machine Learning** demonstrates the strong potential of ML-based systems in transforming academic monitoring and personalized education. Traditional evaluation methods rely heavily on periodic examinations, which provide only limited insights into student learning patterns. In contrast, the proposed ML-based system continuously analyzes academic data and provides actionable predictions that can significantly improve student learning outcomes.

Machine learning models such as Random Forest, Gradient Boosting, and LSTM proved effective in identifying performance trends, forecasting grades, and detecting at-risk students. With over 92% prediction accuracy, the system shows strong reliability, particularly when dealing with large and complex educational datasets. The study confirmed that factors such as attendance, assignment performance, quiz scores, and engagement frequency play a crucial role in academic success.

The system's tracking module enhances transparency in learning by offering students a visual representation of their performance through dashboards and personalized alerts. This feature helps students identify problem areas early and adjust their study strategies accordingly. Educators also gain valuable insights into classroom dynamics, enabling more effective academic intervention processes.

Moreover, the integration of clustering techniques facilitates personalized learning paths by grouping students according to performance behavior. This allows institutions to tailor instructional methods, distribute resources efficiently, and design targeted support programs for at-risk learners.

While the outcomes of this research are highly positive, several challenges remain. Data imbalance, privacy concerns, and model interpretability require ongoing attention. Ensuring fair predictions for all students and maintaining data confidentiality are essential for real-world deployment. Additionally, incorporating reinforcement learning in future versions may further enhance personalized recommendations by enabling the system to adaptively optimize learning pathways.

In conclusion, the proposed machine learning-based student performance system significantly strengthens academic monitoring by enabling continuous performance tracking, early risk detection, and personalized educational support. The research contributes to the growing field of learning analytics and demonstrates how ML technologies can enhance digital education, improve institutional decision-making, and empower learners to achieve better academic success.

REFERENCES

[1] N. Thai-Nghe, Z. Gantner, and L. Schmidt-Thieme, "Improving academic performance prediction by using ensemble methods and learning analytics," *Journal of Educational Data Mining*, vol. 3, no. 1, pp. 1–28, 2011.
[2] R. S. Baker and P. S. Inventado, "Educational data mining and learning analytics," in *Learning Analytics*, Springer, 2014, pp. 61–75.
[3] S. 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.
[4] A. Mueen, A. Zafar, and U. Manzoor, "Modeling and predicting students' academic performance using data mining techniques," *International Journal of Modern Education and Computer Science*, vol. 5, no. 11, pp. 36–42, 2013.
[5] J. L. H. Huan et al., "Student performance prediction using machine learning: A systematic review," *IEEE Access*, vol. 8, pp. 147298–147312, 2020.