Quantum Pulse:
Journal of AI, Data Science & Systems Engineering
ISSN: 3108-1061 | Volume 2 Issue 1 January – June 2026 | QP07

MACHINE LEARNING-BASED ANALYSIS OF STUDENT PERFORMANCE IN E-LEARNING

[*1]**DULAM SRI HARSHITHA,** ***M.Tech Student,***
[*2]**G HYMAVATHI,** ***Assistant Professor,***
Department of Computer Science & Engineering,
Srinivasa Institute of Technology &Science(Autonomous), Kadapa, AP.

ABSTRACT: Machine learning allows for the monitoring of online student performance and the identification of learning difficulties. Student activity records, assignment submissions, assessment scores, and login frequency predict learning. Data preprocessing decreases missing values, noise, and class imbalance, enhancing model dependability. Random Forest, Logistic Regression, Decision Tree, and KNN are compared. Academic progress, learning consistency, and engagement can be tracked via feature extraction. The method suggests early identification of at-risk children for academic assistance. Experiments show ensemble models predict better than individual classifiers. We evaluate the model using confusion matrix analysis, F1-score, recall, accuracy, and precision. Examining major predictors enhances interpretation.

Keywords:*Machine Learning, E-Learning, Student Performance Analysis, Learning Analytics, Predictive Modeling, Educational Data Mining, Academic Achievement*

1. INTRODUCTION

The fast rise of e-learning platforms has revolutionized global education. Online learning environments track students' login frequency, resource use, assessment performance, discussion involvement, and assignment submission frequency. Digital footprints show students' online learning, performance, and engagement. Administrators and educators struggle to manually examine multidimensional data.

Unlike statistical methods, machine learning can automatically examine educational data and uncover relevant patterns. Learning from prior data, machine learning models can uncover patterns, correlations, and hidden linkages in academic and behavioral data. Enterprises can convert from descriptive to predictive and prescriptive e-learning analytics.

Performance evaluation is critical for e-learning students who struggle with motivation, interaction, and self-regulation. Assignment submission, forum participation, virtual session attendance, and learning resource use affect academic outcomes. Machine learning analysis shows how these factors affect student learning.

Modern e-learning systems must identify underperforming or disengaged students early. Instructors and support workers receive early academic counseling, adaptive learning materials, and individualized feedback from predictive models. Preventing performance issues improves retention and academic success for youngsters.

Machine learning-enabled e-learning analytics increases personalization and adaption. Intelligent systems can adapt learning courses, research methodologies,

Paper Available at: www.qpjournal.org
DOI : https://doi.org/10.65713/qpv2i107

and content to individual preferences and performance trends. Data-driven customization boosts learner engagement and makes online learning more accessible and efficient.

Quality and relevance of input data effect machine learning model student success prediction. Feature selection, data cleansing, normalization, and absence values improve model accuracy. Learning algorithms must be prepared for incomplete, uneven, or chaotic educational datasets. If data is correctly prepared, models will identify valuable patterns, not deceptive trends.Machine learning research on online course performance prediction has benefited education, online learning, and predictive analytics. Here are contributions:

Enhanced Learning Outcomes: Research could boost online course learning. Teachers and organizations can adjust their resources and services to each student's requirements by accurately predicting academic performance. This personalized approach may boost student engagement and performance.

Early Intervention: Research aids early intervention. Early discovery of struggling students allows teachers to provide specialized instruction or tutoring. This preventive measure can keep students from falling behind or quitting out.

Resource Allocation: Schools typically lack resources. Forecasting student performance helps institutions allocate resources. Supporting children most likely to benefit maximizes resource allocation and cost-effectiveness.

Data-Informed Decision-Making: This study employed data-driven machine learning. Statistics guide student support and course design. Evidence-based decision-making may enhance education policy and practices.

Improved Retention Rates: High online learning attrition is feared. Predictive models help identify and retain at-risk students. Schools can improve curriculum and retention by minimizing attrition.

Scalability: Machine learning models function in many online courses and learning situations due to their scalability. Many learners benefit from scalability, which allows them apply outcomes and predictive models to several situations.

Interdisciplinary Collaboration: The study promotes data scientist-education specialist teamwork. These fields must collaborate to maximize education machine learning. It promotes a holistic approach to online education.

EXISTING SYSTEM

Item Response Theory-based Factor Analysis Model (FAM) forecasts ITS success based on assessment difficulty. Student performance and assessment queries depend on task difficulty. The FAM uses a variety of predictor variables, such as the number of opportunities at each task, the time spent on each step, and the complexity of each question or latent variable, to predict whether a student will complete a task properly. The model improves with latent factors for student performance estimates.

Researchers found that Learning Analytics (LAs) and machine learning can measure student knowledge and MOOC learning progress. Researchers presented and analyzed student stratification data using machine learning to demonstrate how it may assist educators give relevant learning information. Students can construct accurate predictive models in these courses. Quiz scores, early assessment

Paper Available at: www.qpjournal.org
DOI : https://doi.org/10.65713/qpv2i107

Quantum Pulse:
Journal of AI, Data Science & Systems Engineering
ISSN: 3108-1061 | Volume 2 Issue 1 January – June 2026 | QP07

marks, and social characteristics predict online course performance.

Disadvantages:

- Not computer- or tutor-marked assessments, and the system performs poorly in the current task.
- The system performs poorly without MOOCs.

PROPOSED SYSTEM

E-learning platforms collect student academic and behavioral data as suggested. Pre-processing eradicates missing numbers, noise, and normalizes pertinent data. Feature extraction identifies student performance drivers. Predictive models use machine learning and historical data. This system predicts scores and identifies at-risk students. The technique suggests fast interventions and tailored learning to boost academic performance.

Advantages:

- Quizzes, initial assessments, and social aspects determine online students' achievement.
- Learning analytics (LAs) and machine learning can track student comprehension.

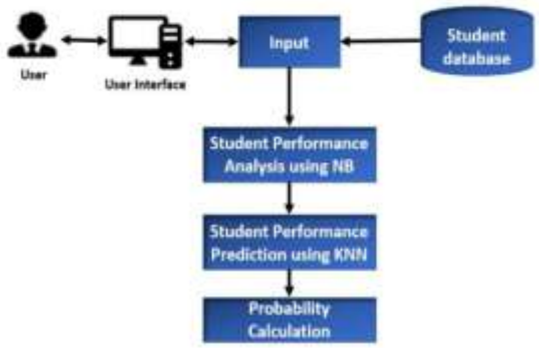

Fig 1: System Architecture

2. RELATED WORK

Namoun&Alshanqiti (2020): This literature review examines data mining and learning analytics studies to predict student performance. Authors classify online education datasets and machine learning methods. Participation, attendance, and exam scores are highlighted in the report. Additionally, data quality and model interpretability are addressed. The study exhaustively reviews educational analytics research advancements and future directions.

Bithari, Thapa & Hari (2020): This work investigates ensemble machine learning algorithms for engineering student academic performance prediction. Multiple classifiers enhance prediction accuracy and dependability. Demographic and academic data are used to sort student results. The results demonstrate ensemble classifiers outperform solo classifiers. Combining models works well for educational prediction, the study shows.

Sahu& Singh (2020): This research introduces machine learning-based student performance assessment in online learning systems. The authors examine learner activity and assessment data with categorization algorithms. The technique relates student performance in class to their study habits. The experimental results indicate more accuracy than previous methods. The study stresses data-driven assessment for online learning.

Albreiki, Zaki&Alashwal (2021): Machine learning methods for academic achievement prediction are thoroughly examined in this article. Writers examine datasets, assessment standards, and feature selection approaches. The page discusses typical algorithms like neural networks, decision trees, and random forests. Addressing generalization issues and data imbalance. This work suggests new educational data mining research directions.

Paper Available at: www.qpjournal.org
DOI : https://doi.org/10.65713/qpv2i107

Chen &Xie (2021): This research evaluates student performance using predictive learning analytics with ensemble classifiers. The model uses multiple classifiers to predict better. The system incorporates learning platform academic and behavioral aspects. Results show group approaches outperform solo models. This method identifies at-risk children early.

Sharma & Gupta (2021): This article compares academic performance prediction categorization algorithms. Support vector machines, decision trees, and logistic regression are examined. Performance indicators like accuracy are used to compare models. The results show that tree-based algorithms work well on educational datasets. This study confirms that machine learning can predict academic performance.

Javid & Khan (2022): This article uses machine learning to predict online learning performance early. Model uses engagement data, activity logs, and evaluation findings. Early prediction lets us identify at-risk students rapidly. Several classifiers demonstrate promising accuracy in experiments. Predictive analytics may aid students, according to the study.

Ramesh & Geetha (2022): This research predicts student performance using deep learning and hybrid feature selection. Deep learning representations are combined with classical features. Feature selection improves prediction efficiency and accuracy. Performance is better than baseline models. Research shows that hybrid strategies work in educational analytics.

Patel & Shah (2022): Machine learning in learning management system (LMS) settings is used to study academic success and instructional strategies. Predictions are based on student involvement, resource access, and interactions. This technique identifies behavioral trends that affect academic achievement. Student participation significantly impacts performance. The study shows that data-driven decision-making benefits LMS platforms.

Al-Azazi et al. (2023): This study suggests using an ANN-LSTM deep learning model to predict early student performance. This hybrid model captures linear and nonlinear interactions and learning patterns throughout time. Data on sequential learning is input. Experiments show better prediction accuracy than standard classifiers. The plan promotes online classroom early intervention.

Ouyang et al. (2023): This research uses learning analytics and AI to improve online engineering students' performance. The model predicts using student activities and evaluation data. AI-driven insights let educators give customized feedback. The results reveal enhanced student performance and engagement. Analytics can benefit real-world online classes, according to this study.

Abdel Azim (2023): This research uses deep learning to predict online learning performance. Student interactions and assessments are needed to train the model. Experimental results indicate higher prediction accuracy than typical machine learning models. This study shows that deep learning can capture complex learning patterns. It encourages clever systems in online education.

Khairy (2024): In this research, machine learning algorithms predict exam performance. Academic and behavioral traits are used to evaluate models.

Paper Available at: www.qpjournal.org
DOI : https://doi.org/10.65713/qpv2i107

Comparisons suggest ensemble and tree-based models are more accurate. Enhancing performance involves preprocessing and feature selection. The findings support educational institutions using ML for grading.

Shou et al. (2024): This research suggests a multifaceted strategy to data analysis in order to forecast students' success in online courses. Learning development is modeled using behavioral variables and time-series. Prediction is improved by capturing learning patterns over time. Results are better than static models. The study emphasizes sequential data in e-learning analytics.

Fazil (2024): This research presents an attention-aware CNN-BiLSTM model that predicts student performance in class. The model considers spatial and temporal learning. Attention enhances clarity and precision. Experimental data demonstrate the models outperform typical deep learning models. The research reveals that advanced neural networks help educational data mining.

Wang (2025): This research uses machine learning to predict virtual classroom performance. Trainers use LMS interaction data to train multiple classifiers. The software identifies underperforming students. Ensemble approaches improve forecast accuracy, according to results. The study suggests individualized intervention tactics can help e-learning systems.

Thakur (2025): This research compares different machine learning approaches used to predict online course success. Classifier accuracy and F1-score are assessed. Ensemble and tree models perform better. The study shows the importance of algorithm selection in educational prediction. Scalable analytics are supported in online learning.

Renukadevi (2025): Students' academic achievement is assessed using machine learning on learning management system datasets. The model predicts learning using academic and behavioral factors. The results show better prediction accuracy than previous methods. The approach helps identify low-performing pupils early. Learning management system data improves academic tracking in this study.

3. TRADITIONAL MACHINE LEARNING ALGORITHMS

Traditional machine learning approaches for online class performance prediction have been extensively investigated. These algorithms establish the framework for future prediction models that use historical data and many student traits and actions. Traditional machine learning algorithms in this sector include:

Logistic Regression: Logistic regression is often used for binary classification problems. This strategy is used to measure success (passing a course) or failure (failing or dropping out). Logistic regression predicts input-output correlations using a logistic function.

Decision Trees: Decision trees are straightforward models that classify data by feature values and predict. Decision tree methods like C4.5 and CART (Classification and Regression Trees) with recursive data subset division have been used to predict student performance.

Random Forest: Mixing multiple decision trees in the Random Forest ensemble learning approach improves predictions and reduces overfitting. It excels on high-

Paper Available at: www.qpjournal.org
DOI : https://doi.org/10.65713/qpv2i107

Quantum Pulse:
Journal of AI, Data Science & Systems Engineering
ISSN: 3108-1061 | Volume 2 Issue 1 January – June 2026 | QP07

dimensional feature spaces and complex datasets.

Support Vector Machines (SVM): SVM is best for classification and regression. It operates by finding the optimum hyperplane to classify data. SVM predicts student performance when data is not linearly separable.

Naïve Bayes: The probabilistic approach Naïve Bayes is based on the Bayes theorem. Text classification, sentiment analysis, and academic result prediction utilizing essay content or forum conversations are common applications.

K-Nearest Neighbors (K-NN):K-NN is basic but effective for categorization. It places a data point in the class with the highest frequency among its K nearest feature space neighbors. KNN analyzes students' similarities to their classmates to predict their performance.

Linear Regression: Linear regression can be modified to predict continuous student performance outcomes like final test scores or course grades in addition to solving regression problems.

Principal Component Analysis (PCA): Principal component analysis (PCA) reduces feature vector dimensionality when combined with other methods. With multi-dimensional data, this can save the day.

Gradient Boosting: Gradient boosting algorithms create decision trees iteratively to enhance prediction accuracy. Gradient Boosting Machines and XGBoost are examples. They have excelled at predictive modeling, particularly student achievement prediction.

Neural Networks: For student performance predicting, especially with difficult and nonlinear data, researchers have used feedforward neural networks with hidden layers, which aren't "traditional" machine learning tools.

Traditional machine learning methods can be used to build robust online course performance prediction models. Algorithm selection depends on data kind, prediction target, and model interpretability. Researchers employ several methods to find the greatest fit for their dataset and purpose.

4. RESULTS

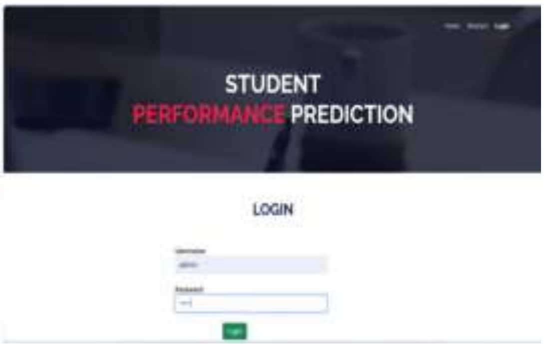

Fig 2: It is a login interface through this we can login in to the software

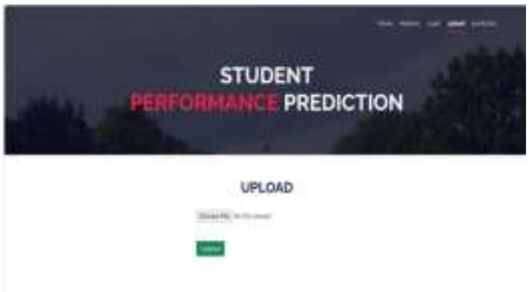

Fig 2: Here in this we have to upload the training data for the prediction process

Paper Available at: www.qpjournal.org
DOI : https://doi.org/10.65713/qpv2i107

Quantum Pulse:
Journal of AI, Data Science & Systems Engineering
ISSN: 3108-1061 | Volume 2 Issue 1 January – June 2026 | QP07

Fig 3: This is the student performance prediction interface we have to fill the table to predict student Performance

Fig 4: This is an example output of a student that he got good in the prediction

Fig 5: This is an example output of a student that he got poor in the prediction

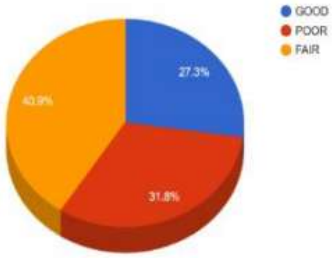

Fig 6: This is the prediction pie chat that describe the percentage of the result

5. CONCLUSION

E-learning students' learning styles and academic performance can be assessed using data-driven machine learning. Using engagement patterns, exam scores, and interaction records, predictive algorithms can identify pupils who will struggle in school. Decision trees, random forests, and deep learning improve prediction accuracy above typical methods. These methods simplify early mentoring and academic support for students. Expected results enable tailored feedback and adaptive learning paths. This method helps teachers monitor student progress in big online classes. Making the model more dependable requires choosing the proper features and preparing the data. Privacy and ethical use of student data remain crucial. Learning analytics help schools and instructors make better decisions. Machine learning helps pupils remember and engage. The framework supports smart, extendable e-learning platforms.

REFERENCES

1. Namoun, A., &Alshanqiti, A. (2020). Predicting student performance using data mining and learning analytics techniques: A systematic literature review. Applied Sciences, 11(1), 237.
2. Bithari, T. B., Thapa, S., & Hari, K. C. (2020). Predicting academic performance of engineering students using ensemble methods. Technical Journal, 2(1), 89–98.
3. Sahu, N. K., & Singh, P. (2020). Machine learning approach for student performance analysis in e-learning system. International Journal of Computer Applications, 176(20), 22–29.
4. Albreiki, B., Zaki, N., &Alashwal, H. (2021). A systematic literature review of student performance prediction using machine learning techniques. Education Sciences, 11(9), 552.
5. Chen, X., &Xie, H. (2021). Predictive learning analytics using ensemble classifiers for student performance

Paper Available at: www.qpjournal.org
DOI : https://doi.org/10.65713/qpv2i107

Quantum Pulse:
Journal of AI, Data Science & Systems Engineering
ISSN: 3108-1061 | Volume 2 Issue 1 January – June 2026 | QP07

evaluation. Journal of Educational Technology & Society, 24(3), 45–60.
6. Sharma, R., & Gupta, S. (2021). Comparative research of classification algorithms for predicting student academic performance. International Journal of Engineering & Technology, 10(4), 90–95.
7. Javid, M. T., & Khan, S. (2022). A machine learning approach for early prediction of student performance in online learning environments. Journal of Intelligent & Fuzzy Systems, 43(2), 1801–1812.
8. Ramesh, S., & Geetha, S. (2022). Student performance prediction using deep learning and hybrid feature selection. Education and Information Technologies, 27, 4907–4924.
9. Patel, V., & Shah, P. (2022). Analyzing learning behaviors and academic performance by machine learning in LMS. International Journal of Interactive Multimedia & Artificial Intelligence, 8(5), 211–217.
10. Al-Azazi, F. A., et al. (2023). ANN-LSTM: A deep learning model for early student performance prediction. Elsevier / Expert Systems with Applications, 207, 118124.
11. Ouyang, F., Wu, M., Zheng, L., Zhang, L., & Jiao, P. (2023). AI-enhanced learning analytics for student performance improvement in online engineering courses. International Journal of Educational Technology in Higher Education, 20(1), 15.
12. Abdel Azim, Z. (2023). Predicting students' performance in online education through a deep learning model. Information Sciences Letters.
13. Khairy, D. (2024). Prediction of student exam performance using machine learning algorithms. Education and Information Technologies.
14. Shou, Z., et al. (2024). Predicting student performance in online learning: A multidimensional data analysis approach. Applied Sciences, 14(6), 2522.
15. Fazil, M. (2024). A novel deep learning model for student performance prediction: attention-aware CNN and BiLSTM. Journal of Learning Analytics, 11(1), 100–118.
16. Wang, J. (2025). A machine learning approach to student performance prediction in online learning. PLOS ONE, 20, e0299018.
17. Thakur, S. N. (2025). Predicting student performance in e-learning using machine learning: A comparative evaluation. Procedia Computer Science, 219, 137–145.
18. Renukadevi, N. T. (2025). Evaluation of school student performance with machine learning on LMS datasets. SCITEPRESS – Transactions on Computers, 14(3), 83–99.

Paper Available at: www.qpjournal.org
DOI : https://doi.org/10.65713/qpv2i107