



# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

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



**Impact Factor: 8.699**

**Volume 14, Issue 5, May 2025**

# Machine Learning-Based Student Risk Assessment: A Comparative Study of Logistic Regression and Gradient Boosting

**Anwar Saeed MM, Mohammad Sadhiq KV, Mohammed Rinas CK, Mr J.Manoj Prabhakar**

U.G. Student, Department of Computer Science and Engineering, Dhaanish Ahmed Institute of Technology,  
Coimbatore, India

U.G. Student, Department of Computer Science and Engineering, Dhaanish Ahmed Institute of Technology,  
Coimbatore, India

U.G. Student, Department of Computer Science and Engineering, Dhaanish Ahmed Institute of Technology,  
Coimbatore, India

Assistant Professor, Department of Computer Science and Engineering, Dhaanish Ahmed Institute of Technology,  
Coimbatore, India

**ABSTRACT:** Predicting student performance is essential for improving educational outcomes and supporting at-risk learners. This study introduces a smart machine learning-based system that analyzes academic, behavioral, and demographic data to forecast whether students are likely to pass or fail. By leveraging features such as attendance, grades, participation, and assignment records, the model identifies performance patterns with high accuracy. In addition to prediction, the system integrates a recommendation engine that suggests personalized interventions. These may include supplementary learning resources, study strategies, or academic support tailored to each student's specific needs. The goal is to enable educators to make proactive, data-driven decisions that enhance learning outcomes and reduce dropout rates.

This dual-purpose system not only identifies students at risk but also offers actionable solutions to improve their academic performance. By combining predictive analytics with personalized recommendations, the platform fosters a more responsive, student-centered approach to education.

**KEYWORDS:** Student Performance Prediction, Educational Data Mining (OULAD), Predictive Modeling, Recommendation System, Machine Learning, At-Risk Students.

## I. INTRODUCTION

The growing complexity of educational environments has intensified the need for intelligent systems that can assist in enhancing student outcomes. With increasing class sizes, diverse student backgrounds, and varied learning paces, educators often face challenges in identifying students who are at risk of academic failure. Traditional methods of monitoring student performance are often reactive and may not provide timely insights needed for effective intervention. This calls for proactive, data-driven approaches that leverage modern technologies to support decision-making in education.

Machine learning has emerged as a powerful tool for educational data mining, capable of analyzing large and diverse datasets to uncover hidden patterns and trends. By utilizing data such as attendance, grades, assignment submissions, and engagement metrics, machine learning models can accurately predict whether a student is likely to pass or fail. Such predictive systems allow educators to intervene early, providing the support needed before students reach a critical point of disengagement or failure.

To further enhance the impact of prediction models, this project incorporates a recommendation system that offers personalized learning strategies and resources based on each student's unique academic profile. These recommendations may include targeted reading materials, tutoring opportunities, or study habit improvements—helping students address specific weaknesses and enhance their performance.

The integration of performance prediction with a recommendation engine creates a comprehensive, intelligent support framework for both students and educators. This approach not only helps identify academic risk early but also empowers learners with actionable insights to improve their outcomes, ultimately contributing to a more adaptive, inclusive, and effective educational system.

## **II. LITERATURE SURVEY**

The use of machine learning in education has gained significant traction, particularly for predicting student performance and enabling early intervention. Romero and Ventura (2007) pioneered educational data mining (EDM), demonstrating how learning management system data can be used to identify learning patterns. Traditional models like Decision Trees, Naive Bayes, and Support Vector Machines have been effectively applied to academic data such as grades, attendance, and assessments (Kotsiantis et al., 2010). Recent studies have expanded these models by including behavioral and engagement metrics, resulting in improved prediction accuracy (Al-Barrak & Al-Razgan, 2016). Deep learning techniques, such as Recurrent Neural Networks, have also been used to model time-dependent student behavior (Piech et al., 2015).

While many works focus on prediction, fewer integrate systems that suggest actionable support. Thai-Nghe et al. (2010) proposed combining prediction with recommendation systems to guide students toward appropriate learning resources. Similarly, Papamitsiou and Economides (2014) emphasized the importance of adaptive learning analytics that not only predict outcomes but respond with personalized interventions. This project builds upon such research by integrating a machine learning-based performance prediction model with a recommendation system, aiming to provide timely, personalized academic support and foster improved educational outcomes.

## **III. METHODOLOGY**

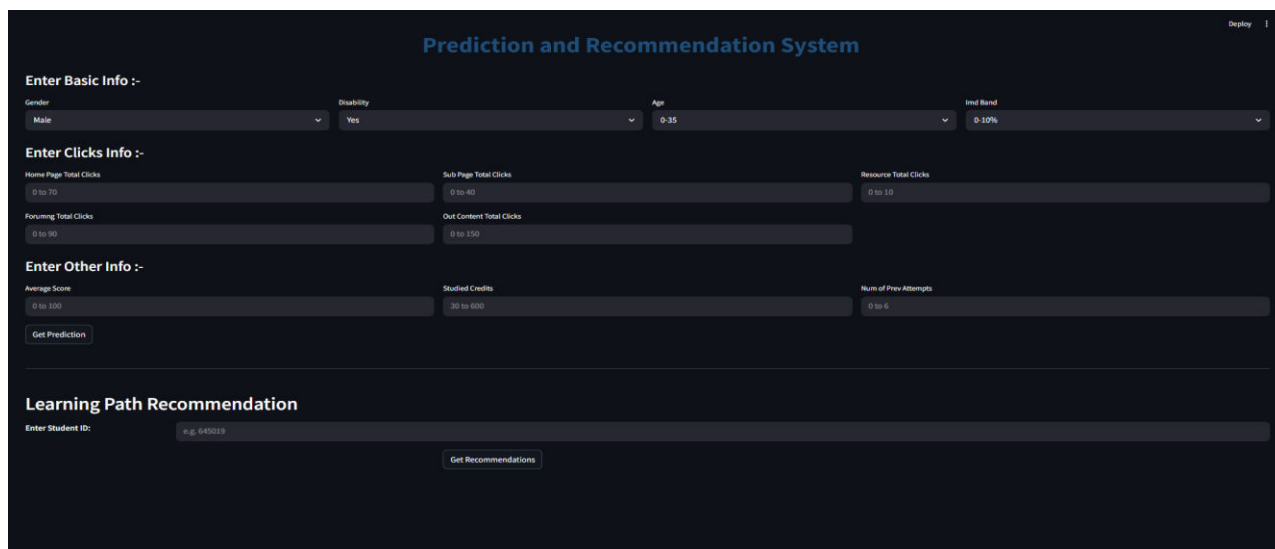
This project adopts a structured approach that combines machine learning for performance prediction with a recommendation system for personalized academic support. The first step involves data collection from sources like OULAD (Open University Learning Analytics Dataset). Key data features include attendance, assignment scores, quiz and exam results, VLE participation, and demographic details. Once gathered, the data undergoes preprocessing, where missing values are handled, categorical variables are encoded, and normalization is applied to ensure consistency and readiness for model training. Following this, feature selection techniques are employed to identify the most influential factors affecting student performance. These selected features are then used to train multiple supervised machine learning models, such as Decision Trees, Logistic Regression, and Gradient Boosting. The models are evaluated using cross-validation and performance metrics like accuracy, precision, recall, and F1-score to determine the most effective predictor.

After prediction, a recommendation system is integrated to provide personalized academic interventions. Based on the prediction outcome and a student's specific weaknesses, the system suggests targeted learning resources, study materials, or tutoring opportunities. Finally, the entire system is deployed in a user-friendly interface that allows educators to monitor performance predictions and implement recommended strategies, ensuring timely support for at-risk students.



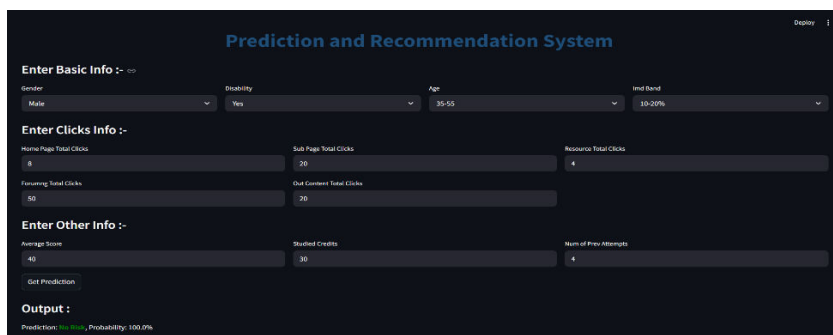
#### IV. EXPERIMENTAL RESULTS

Here is the UserInterface of the Webpage having Prediction and Recommendation system.

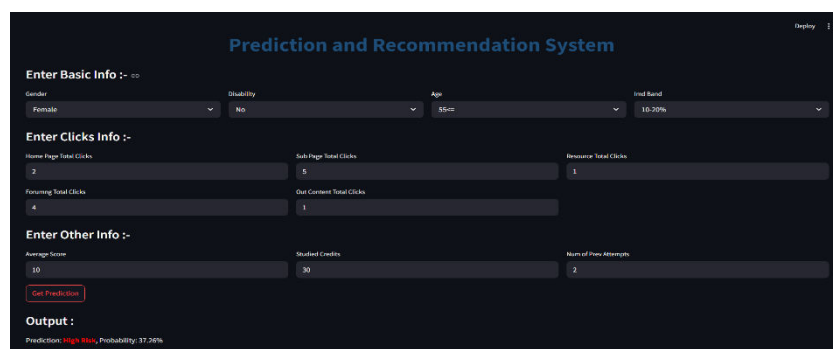


(a)

Fig. 2. After Entering the given details of the student the system will predict the student is in high Risk or not, Image (b) shows the No Risk students and Image (c) shows the High risk students.

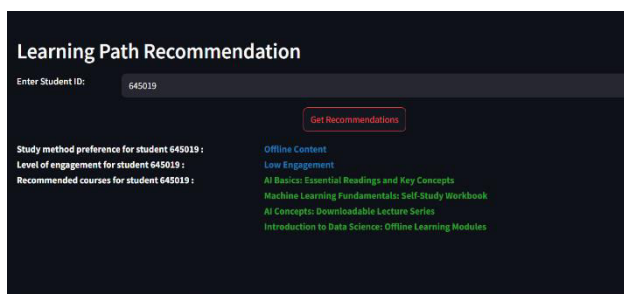


(b)

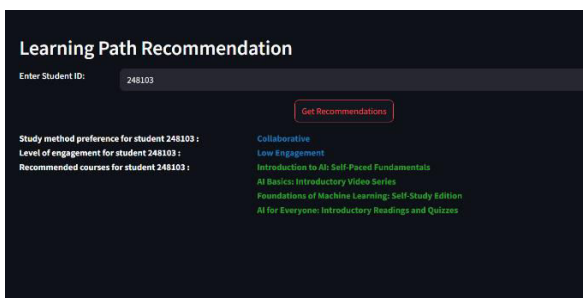


(c)

Fig. 3 Once you enter the students Id which is already in the dataset the recommendation system shows which study approach is best for the student, image(d)&(e) is the results from Recommendation system.



(d)



(e)

## V. CONCLUSION

This project demonstrates the potential of combining machine learning and recommendation systems to enhance educational outcomes. By analyzing academic and VLE data, the system accurately predicts whether a student is likely to pass or fail. More importantly, it goes beyond prediction by offering personalized recommendations tailored to each student's specific needs, such as learning materials or study strategies. This dual approach enables educators to make proactive, informed decisions, allowing timely interventions that can improve performance and reduce dropout rates. Overall, the system supports a more responsive and student-centered learning environment, promoting academic success through data-driven insights and personalized support.

## REFERENCES

- [1] C. Romero and S. Ventura, "Educational data mining: A review of the state of the art," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 40, no. 6, pp. 601–618, Nov. 2010.
- [2] S. B. Kotsiantis, C. J. Pierrakeas, and P. E. Pintelas, "Predicting students' performance in distance learning using machine learning techniques," Applied Artificial Intelligence, vol. 18, no. 5, pp. 411–426, 2004.
- [3] M. A. Al-Barrak and M. Al-Razgan, "Predicting students' performance through classification: A case study," J. Theor. Appl. Inf. Technol., vol. 88, no. 1, pp. 1–8, 2016.
- [4] Z. Papamitsiou and A. A. Economides, "Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence," Educ. Technol. Soc., vol. 17, no. 4, pp. 49–64, 2014.
- [5] N. Thai-Nghe, L. Drumond, T. Horváth, A. Krohn-Grimberghe, and L. Schmidt-Thieme, "Factorization techniques for predicting student performance," in Educational Recommender Systems and Technologies, IGI Global, 2011, pp. 129–153.
- [6] C. Piech et al., "Deep knowledge tracing," in Advances in Neural Information Processing Systems (NeurIPS), vol. 28, pp. 505–513, 2015.
- [7] M. A. Yadav and D. Yadav, "A review of machine learning-based approaches for predicting student performance," Mater. Today Proc., vol. 37, pp. 2780–2786, 2021.
- [8] R. A. Calvo and S. D'Mello, "Affect detection: An interdisciplinary review of models, methods, and their applications," IEEE Trans. Affect. Comput., vol. 1, no. 1, pp. 18–37, Jan. 2010.
- [9] A. M. Tair and R. El-Halees, "Mining educational data to improve students' performance: A case study," Int. J. Inf. Commun. Technol. Res., vol. 2, no. 2, pp. 140–146, 2012.
- [10] M. Asif, A. Merceron, S. A. Ali, and N. G. Haider, "Analyzing undergraduate students' performance using educational data mining," Comput. Educ., vol. 113, pp. 177–194, Oct. 2017.
- [11] E. Ahn, Y. Gil, and E. Putnam-Hornstein, "Predicting youth at high risk of aging out of foster care using machine learning methods," Child Abuse Neglect, vol. 117, Jul. 2021, Art. no. 105059, doi: 10.1016/j.chiabu.2021.105059.
- [12] R. Dembo and W. Walters, "Innovative approaches to identifying and responding to the needs of high risk youth," Subst. Use Misuse, vol. 38, nos. 11–13, pp. 1713–1738, Jan. 2003, doi: 10.1081/ja-120024238.

- [13] H. H. Severson, H. M. Walker, J. Hope-Doolittle, T. R. Kratochwill, and F. M. Gresham, "Proactive, early screening to detect behaviorally at-risk students: Issues, approaches, emerging innovations, and professional practices," *J. School Psychol.*, vol. 45, no. 2, pp. 193–223, Apr. 2007, doi: 10.1016/j.jsp.2006.11.003.
- [14] M. Yağcı, "Educational data mining: Prediction of students' academic performance using machine learning algorithms," *Smart Learn. Environ.*, vol. 9, no. 1, p. 11, Dec. 2022, doi: 10.1186/s40561-022-00192-z.



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details