IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, July 2025
Impact Factor: 7.67

Leveraging Big Educational Data with Machine Learning for Student Success Prediction and Intervention

Kajol Khetan
VP of Marketing, Morgan Publishing House Inc.
kazolkhetan123@gmail.com

Abstract: *An Education provides the tools for success, boosts confidence, and prepares one for life's challenges. Colleges and universities are adapting their pedagogical approaches to take advantage of new technologies, such as artificial intelligence. One of the most important measures of educational success is students' performance in the classroom. A paradigm shift in education has occurred, innovation in technology, especially in the area of AI, has led to this. Applying OULAD, a dataset that includes over 32,000 student records supplemented with demographic information, academic, behavioral, and assessment data, this study delves into the application of machine learning approaches based on artificial intelligence to predict student performance and provide prompt intervention. A comprehensive methodology was employed, beginning with exploratory data analysis through visualizations, followed by data preprocessing, feature selection using the Pearson correlation coefficient, for numerical features, min-max normalization, and one-hot encoding for descriptive variables. The Light Gradient Boosting Machine (LightGBM) was chosen as the best model because of how well it handled big structured datasets compared to the others. The model achieved high predictive accuracy (92.23%), precision (94.40%), recall (93.21%), and F1-score (96.24%), outperforming other models as Logistic Regression, Random Forest, and Support Vector Machines. Results from ROC curves, precision-recall curves, and confusion matrices were used to further confirm the performance, demonstrating the model's robustness and potential to effectively support data-driven educational interventions.*

Keywords: Educational, student success factors, machine learning, deep learning, academic performance prediction, deep learning, educational data analytics, feature selection

DOI: 10.48175/IJARSCT-28469