Predictive Analytics of Student Performance Determinants

Authors : Mahtab Davari, Charles Edward Okon, Somayeh Aghanavesi

Abstract : Every institute of learning is usually interested in the performance of enrolled students. The level of these performances determines the approach an institute of study may adopt in rendering academic services. The focus of this paper is to evaluate students' academic performance in given courses of study using machine learning methods. This study evaluated various supervised machine learning classification algorithms such as Logistic Regression (LR), Support Vector Machine, Random Forest, Decision Tree, K-Nearest Neighbors, Linear Discriminant Analysis, and Quadratic Discriminant Analysis, using selected features to predict study performance. The accuracy, precision, recall, and F1 score obtained from a 5-Fold Cross-Validation were used to determine the best classification algorithm to predict students' performances. SVM (using a linear kernel), LDA, and LR were identified as the best-performing machine learning methods. Also, using the LR model, this study identified students' educational habits such as reading and paying attention in class as strong determinants for a student to have an above-average performance. Other important features include the academic history of the student and work. Demographic factors such as age, gender, high school graduation, etc., had no significant effect on a student's performance.

Keywords : student performance, supervised machine learning, classification, cross-validation, prediction
Conference Title : ICDSML 2022 : International Conference on Data Science and Machine Learning
Conference Location : London, United Kingdom
Conference Dates : October 13-14, 2022