EXPLORING LMS INTERACTION DATA AND PREDICTION OF STUDENTS' GRADES USING STANDARD SUPERVISED CLASSIFICATION ALGORITHMS

Devamullage Pasangi Madhushika Perera

A dissertation submitted in partial fulfilment of the requirement for Master of Science in Business Analytics

School of Computing
Informatics Institute of Technology, Sri Lanka
in collaboration with
Robert Gordon University, UK

2022

Abstract

Higher education providers receive government funding based on graduate completion rather than student enrollment. As a result, incomplete or delayed degree completion is a major source of concern for higher education providers, as these issues have a long-term financial impact and reduce overall cost-effectiveness. As a result, providers must devise strategies to improve educational quality to increase enrollment and retention rates. Predictive modeling techniques are used in this study to help providers identify struggling students in real-time and improve course retention. Prediction models have been developed using student demographic and other behavioral data collected from the Learning Management System to predict students' final grades, allowing lecturers to identify students who are at risk and take proactive steps to help them, such as providing additional support and other timely help. As a result, the findings of this study provide a safety net for both students and academics, while universities can improve student engagement and retention through Learning Management Systems.

Preprocessing techniques and regression methods in machine learning are among the computational methods used in this study. The result is adaptable, and the predictor models are built to work at the level of students. The courses are diverse and complex, and they cannot be generalized. Instead, use a performance measurement method to suggest the best algorithm for predicting students' grades using data from the learning management system. According to research, a learning platform can accurately reflect a student's final grades while also reflecting continuous learning and interaction with the learning management system. This study considers the value of predictive modeling technology with a comprehensive analysis. The online learning platforms are controlled by this model, which can be deduced from independent activities. It establishes a method for conducting an empirical study of predictive modeling using educational data. This research brings learning analytics and the educational process into the realm of mining by demonstrating novel predictive modeling techniques that can be applied to real-world data sets.

By analyzing the LMS log files and students' interactions by activity count in each

activity generated in the years 2020 and 2021, we use machine learning to create models for the early prediction of students' grades in solving learning management system data. Furthermore, because the dataset is drawn from three academic semesters of courses at the University of Sri Jayewardenepura's Faculty of Humanities and Social Sciences, our models are built at the student level. By analyzing learning management system data and detecting students at-risk, failing, and excellent students for all courses offered by the Faculty of Humanities and Social Sciences, University of Sri Jayewardenepura, we can accurately predict students' performance as final grades. For all of the students, we created eight different regression models. Multiple linear regression, Random Fores, and K Nearest Neighbor algorithms are developed and tested as machine learning models, both with and without Principal Component Analysis. The performance of each model was assessed using the R^2 value, the Root Mean Squared Error, and the Mean Absolute Error. As the moment of prediction increases, the R^2 value of all models increases. Without principal component analysis, the Random Forest algorithm performed well, with the highest R^2 value (0.3506) of all the models.

Keywords and terms: students’ grade, student performance, machine learning, linear regression, random forest classification,