Educational big data and learning analytics on predicting students' academic performance

Stephen J.H. Yang, PhD

Chair Professor
Dept. of Computer Science & Information Engineering
National Central University, Taiwan

Abstract:

With the rise of educational big data, learning analytics has become a major trend for improving the quality of teaching and learning. Learning analytics is a kind of data analytics for analyzing students' digital footprint while they interact with learning systems. The objective of this talk is to apply learning analytics on predicting student's academic performance at an early stage and thus provide students with timely intervention. In this talk, I will present how to predict students' academic performance based on tracking log of students' learning activities. We compared the prediction of four datasets from Kyoto University (KU) and National Central University (NCU) with eight classification models. We use the evaluators of accuracy, recall, precision, F1-score, and Area Under the Curve (AUC) of Receiver Operating Characteristic (ROC). According to the prediction results, we found that sample size and feature category have critical influence on the prediction performance. We also found that the significant features based on Pearson correlation analysis have great influence on the prediction performance too. In addition, we address the issues from learning environment, including over-concentration score, dropout students and data instance insufficiently, for improving prediction performance. The results revealed that the proposed performance tuning process could obtain optimal performance metrics and avoid over-fitting problem.

Keywords: educational big data, learning analytics, machine learning, prediction, classification, regression.