

Use of Machine Learning To Assess Factors Affecting Progression, Retention, and Graduation in First-Year Health Professions Students in Qatar: A Longitudinal Study

Dalal Hammoudi Halat ¹, Abdel-Salam G. Abdel-Salam ^{2,3}, Ahmed Bensaid ³, Abderrezzaq Soltani ¹, Lama Alsarraj ¹, Roua Dalli ¹, Ahmed Malki ¹

¹ Academic Quality Department, QU Health, Qatar University, Doha, Qatar

² Department of Mathematics, Statistics, and Physics, College of Arts and Sciences, Qatar University, Doha, Qatar

³ Student Data Management Department, Student Experience Department, Student Affairs, Qatar University, Doha, Qatar

Abstract

Background: Across higher education, student retention, progression, and graduation are considered essential elements of students' academic success. However, there is scarce literature analyzing these attributes across health professions education. The current study aims to explore rates of student retention, progression, and graduation across five colleges of the Health Cluster at Qatar University, and identify predictive factors.

Methods: Secondary longitudinal data for students enrolled at the Health Cluster between 2015 and 2021 were subject to descriptive statistics to obtain retention, progression and graduation rates. The importance of student demographic and academic variables in predicting retention, progression, or graduation was determined by a predictive model using XGBoost, after preparation and feature engineering. A predictive model was constructed, in which weak decision tree models were combined to capture the relationships between the initial predictors and student outcomes. A feature importance score for each predictor was estimated; features that had higher scores were indicative of higher influence on student retention, progression, or graduation.

Results: A total of 88% of the studied cohorts were female Qatari students. The rates of retention and progression across the studied period showed variable distribution, and the majority of students graduated from health colleges within a timeframe of 4-7 years. The first academic year performance, followed by high school GPA, were factors that respectively ranked first and second in importance in predicting retention, progression, and graduation of health majors students. The health college ranked third in importance affecting retention and



8th International Conference on Future of Teaching and Education

Madrid, Spain

26 - 28 April 2024

graduation and fifth regarding progression. The remaining factors including nationality, gender, and whether students were enrolled in a common first year experience for all colleges, had lower predictive importance.

Conclusions: Student retention, progression, and graduation at Qatar University Health Cluster is complex and multifactorial. First year performance and secondary education before college are important in predicting progress in health majors after the first year of university study. Efforts to increase retention, progression, and graduation rates should include academic advising, student support, engagement and communication. Machine learning-based predictive algorithms remain a useful tool that can be precisely leveraged to identify key variables affecting health professions students' performance.

Keywords: Health education, Student progression, Student retention, Student graduation, Machine learning, XGBoost.