

Factors Affecting Grade 11 Students' Academic Performance in Mathematics at a Homagama Zone School

N. D. Kumarasinghe, R. A. B. Abeygunawardana

Department of Statistics, Faculty of Science, University of Colombo, Sri Lanka

Academic performance in mathematics is a key indicator of overall academic success of students and their preparedness for future careers in science, technology, engineering, and mathematics (STEM) disciplines. The objectives of this study are to identify the most important factors that affect academic performance, to identify the relationship between selected variables and academic performance, and to determine the most accurate model for predicting academic performance in mathematics of Grade 11 students. The study was conducted at Vidyarthi National College, Awissawella. The data were collected using a questionnaire. The sampling procedure was a two-stage cluster sampling technique. Primary data were collected through a questionnaire that consisted of 21 five-point Likert scale questions and 18 multiple choice questions. The secondary data were collected from the Department of Examinations, the Zonal Education Office of Homagama, Vidyarthi National College, websites, and research journals. The population size is 463 and the sample size is 209 students. The methodologies that were applied were univariate analysis, bivariate analysis, multiple regression analysis, factor analysis, and random forest classifier. According to the multiple regression analysis, the adjusted R^2 value was 0.669. This means 66.9% of the variability of the third term mathematics score can be explained by the predictors. The academic performance in mathematics of male students was slightly better than that of female students. However, gender is not statistically significant. Considering factor analysis, many student-related factors, two home-related factors, and the occupations of the mother and father were identified as high loading factors. The evaluation was conducted using precision, recall, F1 score, and overall accuracy metrics, and the confusion matrix was used to assess the Random Forest model's performance. The overall prediction accuracy of the model is 93%. According to feature importance, the student-related factors have the highest importance value. Moreover, the occupations of the mother and father, and student related factors also affect academic performance.

Keywords: *Sampling Technique, Multiple Regression Analysis, Factor Analysis, Confusion Matrix, Random Forest Model*