

Predicting Examination Results of Mathematics Using Machine Learning Approach: A Case Study of the G.C.E (O/L) Examination in Sri Lanka

K.T. Nallaperuma* and P.P.G.D. Asanka

Department of Industrial Management, Faculty of Science, University of Kelaniya, Dalugama, Sri Lanka.

*k.t.nallaperuma96@gmail.com

In the Sri Lankan context, government schools provide free education for the students from the primary stage to the collegiate stage. There are 627,477 pupils enrolled in government schools' senior secondary lower level (Grades 10–11). The first bar examination is the General Certificate of Education (Ordinary Level) examination for those students, and if a student fails to complete it, they won't be able to go ahead with the General Certificate of Education (Advanced Level) examination. The success of secondary education is measured by conducting the O/L examination. Out of all other subjects, Mathematics is one of the main and most important subjects in the O/L examination. Therefore, to get a higher grade in Mathematics, students work hard under lots of pressure. This paper uses students' data to identify key factors that affect students' mathematics results in the O/L examination and explore a model to predict Mathematics results of the O/L examination based on their term test Mathematics marks in Grades 10 and 11. Five hundred students from six schools including 1AB, 1C, and Type 2 categories in the Matara district were chosen to collect data using a questionnaire. Key factors that affect students' mathematics results were identified using the ANOVA test and T-test. The Random Forest Classifier model used only students' term test Mathematics marks and the Gradient Boosting model used term test Mathematics marks and students' social and geographical factors to predict the Mathematics results. The results of students who did the O/L examination earlier were used to measure the accuracy of the predictive model. As a result of this study, teachers and parents can focus more on students that are expected to have low grades for Mathematics, found using the model. Moreover, teachers can identify the factors affecting the grades of a student.

Keywords: Academic Performance Predicting, ANOVA Test, Gradient Boosting, Machine Learning, Random Forest