TEST CASE PRIORITY PREDICTION USING TEXT ANALYTICS AND MACHINE LEARNING ALGORITHMS

Wanduramba Hewage Kinkini Indunil Erangi

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

Department of Business

Informatics Institute of Technology, Sri Lanka
in collaboration with
Robert Gordon University, Aberdeen

2022

Abstract

This research focuses on test case prioritization using keyword similarities in the test cases. If there are two test cases with similar words, they would put into the same category of priority. High prioritized test cases are often used for smoke and sanity testing after modifications of functionalities. High and Medium prioritized test cases always selected for Regression testing. This is suitable for ongoing and long going projects with frequent team member such as a previous test engineer was removed from the project and assigned someone else to the team, it becomes a difficult task for the new member to catch up the work. Maintaining large number of regression test suites and other test cases is crucial when customer satisfaction is on the plate. Classification algorithms are built, and new test scenario prioritization will be predicted. The results show how useful the proposed model is maintaining test suites in a large, critical project.

When manually decide test priority level, we must have a proper knowledge on domain, some special keywords relate to the project and deep clarification on client requirement. In this research, multiple algorithms were applied on same project dataset and evaluated about their ability to categorize test cases into 3 main categories.

I introduce a test case prioritization method based on keywords of the previous test cases related to a specific project. 2416 test cases were trained under 15 different machine learning algorithms and could get 98.96% accuracy. This high number of accuracy level indicates that based on keywords used in test cases, we can differentiate the priority level of each.