INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER

Test Case Selection and Prioritization Using Machine Learning

A dissertation by

Mr Sudeeptha Goonewardene

Supervised by

Mr Prathieshna Vekneswaran

Submitted in partial fulfilment of the requirements for the BSc in Computer Science degree at the University of Westminster.

May 2023

Abstract

Techniques for selecting test cases shorten testing duration and cost, but they could omit some crucial tests that can detect vulnerabilities. Test case prioritization considers all test cases and performs them up until resources are depleted or all test cases have been executed, while continuously focusing on the most important ones. Over time, machine learning has gained popularity in software engineering and is now utilized to solve many different problems. Learning concerns can include issues with software development and maintenance, and machine learning techniques have been very effective in resolving these problems. Machine learning approaches have been used to solve the test case prioritization problem, extending its range of use.

This project will involve conducting research on test case prioritization using machine learning techniques that have been considered in the past. Additionally, the project's primary focus will be on the gaps and problems in earlier studies in the relevant field. This project proposal emphasizes the problem that will be addressed, the novelty of the research idea, and the research gap by using the pertinent evidence from previous research evaluations. At the conclusion of this research, methodology and delivery plans are also covered.

A supervised learning approach with a few machine learning algorithms was used for test case priority classification. Logistic Regression, SVM, Random Forest, Decision Tree, and KNN were tested separately to evaluate the performance of each model. To enhance the accuracy of the final classification model, an ensemble learning technique was used. From the ensemble method, a Voting classifier was used with the above-mentioned ML algorithms to build the final test case priority classifier.

Keywords: Test case prioritization, Machine learning, Supervised learning, Ensemble learning, Voting classifier