

DELIVERY CAPABILITY PREDICTION IN ITERATIVE SOFTWARE DEVELOPMENT DURING MID SPRINT

Shareefa Mareyam Mohamed

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

Department of Business

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

2020



SCHOOL OF COMPUTING SCIENCE AND DIGITAL MEDIA

MSc Business Analytics

Abstract

Agile methodology or framework which is an iterative software development method has increased the rate of success of a software project in most of the companies tremendously over the past few years. Fast delivery, on-time delivery, quality product and incremental delivery on an iteration basis, has become a must in the modern, trending software world today. As an iteration or sprint is on-going, forecasting and observing the progress of the iteration and understanding the delivery capability is essential.

Proper monitoring will in return provide quality software applications, while helping to save the extra cost and complete the project within the allocated budget. This project is an approach to support and help understand the project managers, project stakeholders, product owners, other interested parties and decision-makers about the capability or possibility of delivering an ongoing project iteration. The approach includes a set of historical project-related data that has been extracted in forms of reports and combined with statistical techniques and machine learning techniques to predict the delivery capability.

This project proposes two models and the 1st model is to predict the project delivery capability and the 2nd model to predict the possible spillover story points. The selected models and the ideas were discussed with experts and professional to get an understanding and the same was discussed to identify the key factors which affect the delivery capability. Two sources of reports were identified which can have an influence over the delivery capability, those are the project sprint wise detailed report and the respective resource-related report.

Supervised learning algorithms in the contexts of classification and regression was used. Finally, the defined models were visually presented for future prediction purposes along with an interactive dashboard.

Based on the study it can be evaluated that the attributes: sprint duration, sprint start story points, extra story points added after the sprint start, removed story points after

the sprint start, completed story points or completed work by the mid of a sprint and dev members influence a sprint delivery capability.

The first model that is the logistic regression model reached an accuracy level of 95% and the second model that is the support vector regression model reached an accuracy level of 87%.