Automate software requirements classification using machine learning algorithms

By

Gaith Yaser AL Ali

Supervisor

Dr. Ahmad Althunibat

Al-Zaytoonah University of Jordan, 2021

Abstract

Requirements software engineering is the process of defining, documenting, and maintaining requirements in the engineering design process. It is a common role in system and software engineering. Requirements are categorized into two main categories, namely, functional and non-functional requirements. In large software, manual classification of requirements requires a lot of effort, time, money, and results in manual classification are usually less accurate. When get more correct the classification of requirements, then become more accurate the software . To solve this problem, this work seeks to find a solution for requirements classification by automation using machine learning by text feature extraction and machine learning algorithms, and for the process of automatic classification, these texts are cleaned using a series of normalization steps, then applying feature extraction and selection techniques, then using three machine learning algorithms for classification. When using Term Frequency, and Inverse Document Frequency followed by Support Vector Machine in binary classification, this extracted an improved classification of requirements, with the F scale being 0.92. In multi classifications, Term Frequency, and Inverse Document Frequency and Bag of Word were used, followed by

Support Vector Machine and Logistic Regression algorithms, the results showed an improvement in the classification of new requirements, which appeared on a scale of f-measure 0.99.

Keywords: functional requirements; machine learning; non-functional requirements; requirements; text normalization.