

" تقدير أمثل للجهد القائم على الحالة لمشاريع البرمجيات باستخدام خوارزمية الجينات "

“An Optimized Case Based Software Project Effort Estimation using Genetic Algorithm

Author:

Shaima Hameed Odeh Abusheiban

Supervisor:

Dr. Yousef Elsheikh

Co-Supervisor:

Prof. Mohammed AlazzeH

Master Thesis

**This thesis was submitted in Partial Fulfilment of the Requirements
For Master Degree in Computer Science.**

Deanship of Scientific Research and Graduate Studies

Applied Science Private University

Amman-Jordan

August 2021

Abstract

Software development organizations have always suffered from misestimated efforts for Software projects that lead to tremendous financial, time, and resource losses. Therefore, the research community developed various effort estimation methods hoping to achieve high levels of accuracy and efficiency. One widely popular machine learning (ML) method used with software effort estimation (SEE) is case-based reasoning (CBR) or analogy-based estimation (ABE). The basic concept of CBR is similar to the expert estimation mindset but with an automated process where the final estimate is developed by reusing similar historical projects. Many studies demonstrated that CBR is one of the most accurate methods for prediction and estimation, especially software effort estimation. However, the primary problem of using CBR is its many options or parameters that are very hard to adapt. Therefore, the adaptation process is an essential part of CBR to produce predictions for new projects and an efficient estimation based on tuning the solutions selected using adaptation strategy. Thus, in this study, we use the Genetic algorithm (GA) to find the best combination of parameters to produce the most accurate effort estimations when using CBR to estimate software effort. GA will adapt CBR parameters such as feature weighting and the number of close projects selected. This solution produced significantly accurate results in some of the datasets we have used and reduced lowquality products by reducing software financial, time and resource losses due to issues resulted from effort misestimation. We have experimented our CBR-GA model with various datasets including Albrecht, Maxwell, NASA, Telecom, Kemerer, China, and desharnais and the results obtained diversified from significant improvement to no improvements at all.14.

Keywords: Software effort estimation (SEE), Case-based reasoning (CBR), Genetic Algorithm (GA), Machine learning (ML).