

A Generic Data Mining Model for Software Cost Estimation Based on Novel Input Selection Procedure

Zahid Hussain Wani, University of Kashmir, Srinagar, India

Kaiser J. Giri, Islamic University of Science & Technology, Awantipora, India

Rumaan Bashir, Islamic University of Science & Technology, Awantipora, India

ABSTRACT

It is always preferable for any estimation model to be inclusive as accuracy in estimation models inherently lie with their inclusiveness. Software cost estimation is the prediction of development effort and time required to develop a software project and being predictive in nature, it demands for inclusiveness, which will accordingly bring the accuracy in it. In this study, a generic model for software cost estimation using an input selection procedure is proposed. The proposed model brings inclusiveness into the already available data mining techniques of software cost estimation by sensitively choosing a subset of highly relevant project attributes and ignoring the less relevant ones. In this article, a diverse set of data mining techniques for software cost estimation are considered. All these techniques are experimented on five data sets before and after passed through the proposed procedure. The obtained results showed that newly generated techniques after being passed through the proposed procedure offer accurate results up in the way of efficiency in software cost estimation.

KEYWORDS:

Artificial Neural Network, Data Mining Model, Functional Link Artificial Neural Network, Genetic Algorithm, Input Selection Procedure, Software Cost Estimation

INTRODUCTION

Software cost estimation is the summation of building effort and calendar time required to develop any software project. The building effort includes the count of working hours and the number of workers in terms of staffing levels required to develop a software project. The software development organizations often face the problems of estimations of effort and development time in software development process. A big reason for this is the elusive character of the “software”, software as a product, (Chaos Report, 2009). Most often in the process of software effort estimation, the effort needed to develop any new software project is estimated by comparing this project based on relevance of its attributes like number of Lines of code, development platform used and the developmental team experience with previous projects and accordingly where the current project’s data fits best, the information of that very existing project is used by management for estimation of current software project. This way of following the estimation process of any software project at an individual level lets every single project manager to effectively evaluate his project progress, gave him a potential cost control, delivery accuracy, and at the management level or more precisely in a broader perspective helps the organization in improving the planning and utilization of personnel, making more accurate

DOI: 10.4018/IJIRR.2019010102

Copyright © 2019, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.