Comparative analysis of J48 with statistical and machine learning methods in predicting fault-prone classes using object-oriented systems

Yogesh Singh*
Arvinder Kaur†
Ruchika Malhotra§

Department of Software Engineering
Delhi Technological University
(formerly known as Delhi College of Engineering)
Bawana, Delhi 110403
India

Abstract

There are available metrics for predicting fault prone classes, which may help software organizations for planning and performing testing activities. This may be possible due to proper allocation of resources on fault prone parts of the design and code of the software. Hence, importance and usefulness of such metrics is understandable, but empirical validation of these metrics is always a great challenge. J48 decision tree algorithm has been successfully applied for solving classification problems in many applications. This paper evaluates the capability of algorithm and compares its performance with nine statistical and machine learning methods in predicting fault prone software classes using publicly available NASA data set. The results indicate that the prediction performance of J48 is generally better than other statistical and machine learning models. However, similar types of studies are required to be carried out in order to establish the acceptability of the J48 model.

Keywords: Software metrics, Object-oriented, fault prediction, machine learning, decision tree, software quality