Abstract

Recently, early prediction of software defects using the machine learning techniques has attracted more attention of researchers due to its importance in producing a successful software. On the other side, it reduces the cost of software development and facilitates procedures to identify the reasons for determining the percentage of defect-prone software in future. There is no conclusive evidence for specific types of machine learning that will be more efficient and accurate to predict of software defects. However, some of the previous related work proposes the ensemble learning techniques as a more accurate alternative. This paper introduces the resample technique with three types of ensemble learners; Boosting, Bagging and Rotation Forest, using eight of base learner tested on seven types of benchmark datasets provided in the PROMISE repository. Results indicate that accuracy has been improved using ensemble techniques more than single leaners especially in conjunction with Rotation Forest with the resample technique in most of the algorithms used in the experimental results.

Index Terms
Computer Science
Software Engineering
Keywords

Software Defects, Ensemble methods, Resample Technique, Base Learner, Bagging, Boosting, Rotation Forest.