International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 09 | Sep 2020 www.irjet.net p-ISSN: 2395-0072

META ANALYSIS ON CROSS PROJECT DEFECT PREDICTION

B. Santhi Priya[1], D. Sowmya[2], D. Vijaya Krithika[3], M. Yuvashree[4S]

[1]Assistant professor. of Information Technology, Jeppiaar SRR Engineering College, Chennai.
[2,3,4]Final Year Student, Dept. of Information Technology, Jeppiaar SRR Engineering College, Chennai.

Abstract :- Tender sometimes refers to the method whereby governments invite bids for big comes that has got to be submitted inside a finite point. Tendering is that the method of constructing a proposal, bid or proposal, or expressing interest in response to letter of invitation or request for tender A young may be a submission created by a contractor in response to letter of invitation to tender .It makes another for the of products or services. Tender documents is also ready for a variety of contracts, like instrumentality provide, the main construction contract, demolition, facultative works, etc. This study proposes a machine-learning-inspired approach which will be wont to predict the defects in an exceedingly efficient manner. The end result of such a prediction will facilitate to assess that provider is probably going to supply best worth for a selected project. This paper summarizes varied methodologies employed by various authors. Models supported Nearest-Neighbour and call Tree tends to perform well in CPDP. Once a typical approach is applied for defect prediction, promising results will be obtained through continuous refinement of the prediction model North American nation

1. INTRODUCTION

Cross Project Defect Prediction (CPDP) recently gained considerable attention, yet there are no systematic efforts to analyze existing empirical evidence. The objective of this project is to summaries, analyze and assess the empirical evidence regarding metrics, modeling techniques, different approaches and performance evaluation criteria in the context of CPDP. Different modeling techniques (Machine Learning and Regression) are used in the context of CPDP. Logistic Regression (LR), Naïve Bayes(NB), Bayesian Networks (BN), Nearest-Neighbor (NN), Decision Tree (DTree) and Random Forest (RF) are some of the popular techniques. The goal here is to determine which modeling techniques are frequently used in CPDP and how they relatively perform. The principle of defect prediction is to learn a model from a corpus of data and apply the model to new and unseen data. The training data can be from the same project, i.e., within project defect prediction (WPDP) or from other projects, i.e., cross project defect prediction (CPDP). The goals of defect prediction are to identify the faulty units of code, to estimate the defects in the system, to track and locate faulty changes, classes, functions or statements, for optimal use of the available quality assurance resources. We can use this CPDP in Examination halls for analyzing the question paper. Further, we want to explore the relative performance of CPDP vs. Within Project Defect Prediction (WPDP) models. CPDP can be further used in much kind of applications for detecting the defects in the applications by analyzing them

2. RELATED WORKS

Faimison Porto, Leandro Minku, Emilia Mendes, Adenilso Simao **[1]**The prediction of defects in a target project based on data from external projects is called Cross-Project Defect Prediction (CPDP). Several methods have been proposed to improve the predictive performance of CPDP models. However, there is a lack of comparison among state-of-the-art methods. Moreover, previous work has shown that the most suitable method for a project can vary according to the project being predicted. This makes the choice of which method to use difficult. We provide an extensive experimental comparison of 31 CPDP methods derived from state-of-the-art approaches, applied to 47 versions of 15 open source software projects. Four methods stood out as presenting the best performances across datasets. However, the most suitable among these methods still varies according to the project being predicted. Therefore, we propose and evaluate a meta-learning solution designed to automatically select and recommend the most suitable CPDP method for a project. Our results show that the meta-learning solution is able to learn from previous experiences and recommend suitable methods dynamically. When compared to the base methods, however, the proposed solution presented minor difference of performance. These results provide valuable knowledge about the possibilities and limitations of a meta-learning solution applied for CPDP.

Master's Thesis, Dimuthu Gunarathna [2] The majority of the Cross Project Defect Prediction (CPDP) models have been constructed using combinations of different types of independent variables. The models that perform well tend to be using combinations of different types of independent variables. Models based on Nearest Neighbor (NN) and Decision Tree (DTree) appear to perform well in CPDP context. Most commonly used Naive Bayes (NB) seemed to having average performance among other modeling techniques. Recall, precision, F-measure, probability of false alarm (pf) and Area Under Curve (AUC) are the commonly used performance metrics in cross-project context. Filtering and data transformation are also frequently used approaches in the cross-project context. The majority of the CPDP approaches address one or more data related issues using various row and column processing methods. Models appear to be performing well