

META ANALYSIS ON CROSS PROJECT DEFECT PREDICTION

B. Santhi Priya¹, D. Sowmya², D. Vijaya Krithika³, M. Yuvasree^{4S}

¹Assistant 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

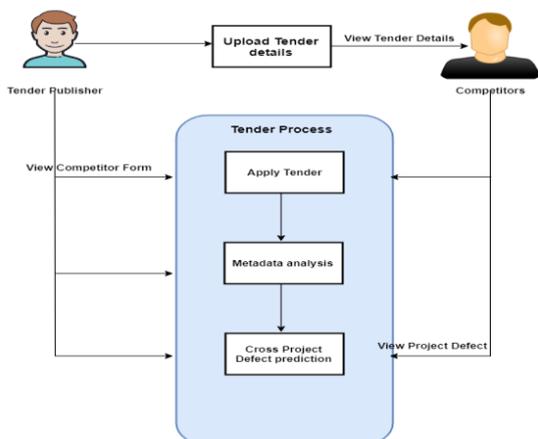
when filtering approach is used and model is built based on NB. Further, models perform well with data transformation approach is used and model is built based on Support Vector Machine (SVM). There is no significant difference in performance of CPDP models compared with Within Project Defect Prediction (WPDP) model performance. CPDP model perform well in majority cases in terms of recall.

Marian Jureczko, Lech Madeyski [3] The paper presents an analysis of 83 versions of industrial, open-source and academic projects. We have empirically evaluated whether those project types constitute separate classes of projects with regard to defect prediction. Statistical tests proved that there exist significant differences between the models trained on the aforementioned project classes. This work makes the next step towards cross-project reusability of defect prediction models and facilitates their adoption, which has been very limited so far.

3. PROPOSED SYSTEM

We used an integrated machine learning approach based on regression models constructed using a set of predictor variables. Based on these variables, our regression models estimate the defects in the product prior to testing through multiple and simple linear regression techniques. Our approach is distinct from previously proposed prediction models. The key premise of CPDP is to learn from data from a set of projects and then to apply resulting models to another set of projects. Therefore, in presence of relevant data from other projects, CPDP has a great practical potential. The prediction outcomes with data from other projects can be of great value as even a tiny decrease in the bug rates can lead to significant financial savings in terms of quality assurance costs, as opposed to exponential growth in repair costs and damages, as a result of failure to discover bugs in a timely manner. Thus, it is demonstrated that this technique can provide a blueprint for overture testing to enhance the effectiveness of tender evaluation activities.

4. PROPOSED ARCHITECTURE



5. MODULES AND DESCR IPTION

5.1. ADMIN AND USER

REGISTER:

View and Authorize Users:

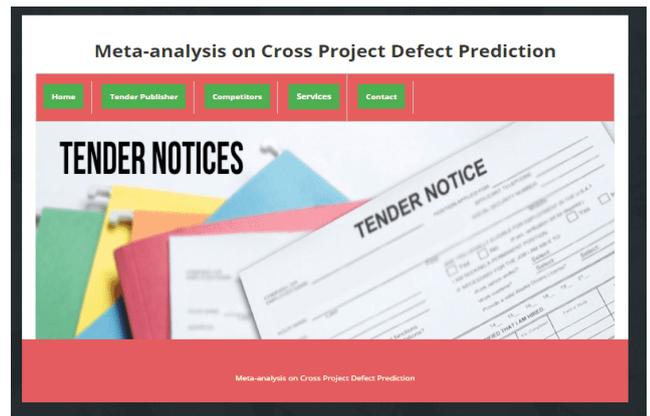
In this module, the admin can view the list of users who all registered. In this, the admin can view the user’s details such as, user name, email, address and admin authorizes the users.

User:

In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like View Profile, Add Category, and apply tender.

LOGIN:

The Login Module is a portal module that allows users to type a user name and password to log in. This module can be placed on any module tab to allow users to log in to the system. The module is no longer available to users after they have logged in.



5.2 TENDER PUBLISHER

Tender usually refers to the process whereby governments invite bids for large projects that must be submitted within a finite deadline. Tendering is the process of making an offer, bid or proposal, or expressing interest in response to an invitation or request for tender. From the time the tender is released, it is generally in the market for 25 calendar days. Pre-release preparation, the advertisement period, evaluations, Tender Board endorsement and agreement finalization generally result in a tender process taking an average of 90 days to complete. A tender is a submission made by a contractor in response to an invitation to tender. It makes an offer for the supply of goods or services. Tender documents are prepared to seek offers Tender documents may be prepared for a range of contracts, such

as equipmentsupply, the mainconstructioncontract (including design bythe contractor), demolition, enabling works, etc.

DOMESTIC TENDER:

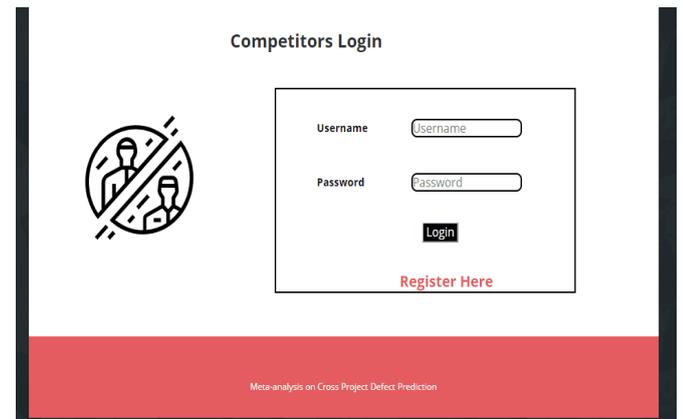
The Tender Which Is Circulate Or Invoked Inside Nation Is Called A Domestic Tender. A Tender Is A Submission Made By A Prospective Supplier In Response To An Invitation To Tender. It Makes An Offer For The Supply Of Goods Or Services. In Construction, The Main Tender Process Is Generally For The Selection Of The Contractor That Will Construct The Works.

INTERNATIONAL TENDER:

Tenders Info is a online government Tender information Provider Company, helping business across globe in finding business opportunities. Our Tender Portal provide information on E-procurement, Public Tenders, Government bids, International Bidding opportunities etc. We help business from all sectors and countries with potential tender leads that help them grow their business globally through Tendering. An international tender is a competitive selection process carried out by a public body to contract a service, carry out a work or acquire goods. International tenders are open to any company since the important thing is not the size, but the economic situation.

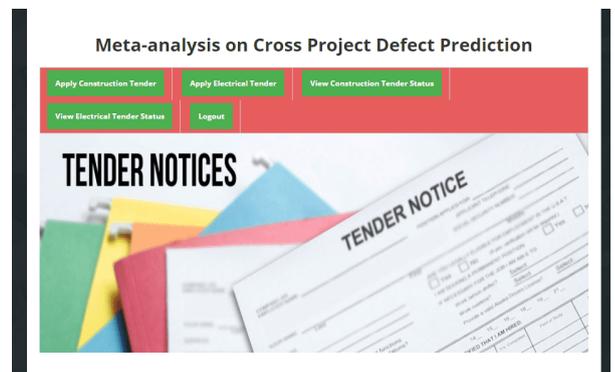
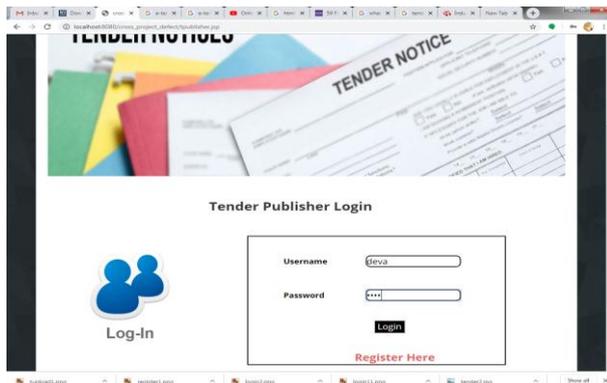
using authorized user name and password. Once Login is successful user will do some operations like View Profile, Add Category, and apply tender.

The Login Module is a portal module that allows users to type a user name and password to log in. This module can be placed on any module tab to allow users to log in to the system. The module is no longer available to users after they have logged in.



5.4 COMPETITORS ANALYSIS FORM

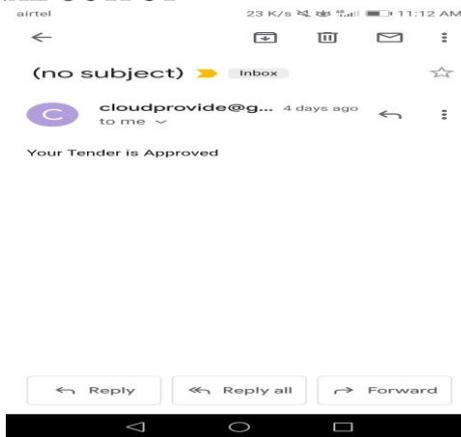
Any person or entity which is a rival against another. In business, a company in the same industry or a similar industry which offers a similar product or service. The presence of one or more competitors can reduce the prices of goods and services as the companies attempt to gain a larger market share. Any person or entity that is in the same industry or a similar industry, or which offers a product similar to that of another person or entity, is called its competitor. These two tests help analysis a market with respect to its attractiveness to competitors.



5.3 COMPETITORS

In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by

5.5 FINAL OUTPUT



hybrid-based approaches," IEEE Tran. on Sys., Man, and Cyber., Part C (Applications and Reviews), vol. 42, no. 4, pp. 463–484, July 2012. [Online]. Available: <http://dx.doi.org/10.1109/TSMCC.2011.2161285>

[5] S. Prateek, A. Pasala, and L. Moreno Aracena, "Evaluating performance of network metrics for bug prediction in software," in Asia-Pacific Soft Eng Conf (APSEC2013), vol. 1, 2013, pp. 124–131. [Online]. Available: <http://dx.doi.org/10.1109/APSEC.2013.27>

[6] F. Zhang, Q. Zheng, Y. Zou, and A. E. Hassan, "Cross-project defect prediction using a connectivity-based unsupervised classifier," in Proc Int Conf on Soft Eng (ICSE2016). New York, NY, USA: ACM, 2016, pp. 309–320. [Online]. Available: <http://dx.doi.org/10.1145/2884781.2884839>

[7] J. C. Platt, "Fast training of support vector machines using sequential minimal optimization," in Advances in Kernel Methods. MIT Press, 1999, pp. 185–208

6. CONCLUSION

This method is termed Cross-Project Defect Prediction. The main purpose of CPDP is to predict defect-prone instances in a project based on the defect data collected from other projects on those public software repositories like promise. In this application the cross project defect prediction is used to predict the defects present in the project and sends the report to the project admin. In upcoming technologies, this concept can be used for the analysis of the defect present in the system and also to send notifications to the right people when a bug needs to be fixed, tested, or marked as resolved

7. FUTURE ENHANCEMENT

This method is termed Cross-Project Defect Prediction (CPDP). The main purpose of CPDP is to predict defect-prone instances (such as classes) in a project based on the defect data collected from other projects on those public software repositories like promise. In this application the cross project defect prediction is used to predict the defects present in the project and sends the report to the project admin. But this process require more time .In future we should overcome these application more efficient and effectively

8. REFERENCES

[1] A Systematic Study of Cross-Project Defect Prediction with Meta-Learning Faimison Porto, Leandro Minku, Emilia Mendes, Adenilso Simao1802.06025v2 [cs.SE] 31 May 2019.

[2] A Systematic Literature Review on Cross-project Defect Prediction University of Oulu Information Processing Science Master's Thesis Dimuthu Gunarathna21.10.2016.

[3] Cross-Project Defect Prediction with Respect to Code Ownership Model: an Empirical Study. e-Informatica Software Engineering Journal, Volume 9, Issue 1, 2015, pages: 21–35, DOI 10.5277/e-Inf150102.

[4] M. Galar, A. Fernandez, E. Barrenechea, H. Bustince, and F. Herrera, "A review on ensembles for the class imbalance problem: Bagging-, boosting-, and