Classification of incident tickets: xxxxx

January 28, 2019

1 Introduction

In development industries (software or hardware) unexpected behavior in the products, called as bugs or issues, are inevitable. The bug reports and failure reports of product are documented in incident tickets and these incident tickets are managed via ticket tracking systems [2] like Bugzilla, GNATS, Jira etc. The tracking systems allow both developers and clients to:

- Post problems encountered with the product

- Suggest enhancements or solutions

- Comment on reports and suggested solution.

Every incident ticket is examined and assigned to the product component in which the problem arose. In many of the industries, this is done manually. Studies [3, 4, 6] suggest that manual classification of bugs leads to miss-classification problem, ie., the tickets are reported to wrong components.

This poor quality of ticket classification had become a serious problem in Bosch. In Bosch tickets are reported in two levels

1. Tickets from the Internal testers.

2. Tickets form the clients.

Technical Manger distributes these tickets to the developers. Developers are the people who are responsible to fix these tickets. This distribution is made based on their expertise and availability. A developer spends 5-10 minutes to get a brief idea about the ticket and spends around one day to study the ticket completely. On an average it takes at least two days to solve an issue. But when the developer gets a ticket that is wrongly assigned or classified all the effort is wasted. This wrongly assigned ticket has to be reassigned to other component based on the developer's knowledge. As per the sources around 2,000 Euros are spent on solving a ticket. These cases are leading to huge resource loss.

We can frame this problem as a Multi-Class Classification problem, by making the component as the class. At the end of the day the challenge is to assign the ticket to the correct component. Each ticket contain fields like Ticket ID, Assigned date, Component name, Summary, Description etc. By using the Fields Summary and Description, a classification model can be built.

2 Related Work

In [5] Johan Anvik et.al has used Support vector Machines, Navie Bayes and C4.5 machine learning algorithms to classify bugs in open reporsitories like Bugzilla. N.Pingclasai et.al in [7] has proposed a model using Topic modeling to classify bugs from other requests. In his work, Latent Dirichlet Allocation (LDA) is used for topic modeling and machine learning algorithms like Decision trees, Navie Bayes classifier and Logistic Regression were used as models.G. Antoniol et.al [4] has proposed a models using Alternating Decision Trees, Naive Bayes classifier

and logistic regression to perform classification on issues in Bug Tracking System. In [1] Ayan Nigam has proposed a self training algorithm by using semi supervised supper vector machine for multi class classification.

3 Proposed Approach

References

[1] A. Nigam, B. Nigam, C. Bhaisare and N. Arya, "Classifying the bugs using multi-class semi supervised support vector machine," International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME-2012), Salem, Tamilnadu, 2012, pp. 393-397. doi: 10.1109/ICPRIME.2012.6208378

[2] D. Kim, Y. Tao, S. Kim and A. Zeller, "Where Should We Fix This Bug? A Two-Phase Recommendation Model," in IEEE Transactions on Software Engineering, vol. 39, no. 11, pp. 1597-1610, Nov. 2013. doi: 10.1109/TSE.2013.24

[3] F. Thung, S. Wang, D. Lo, and L. Jiang, An empirical study of bugs in machine learning systems, in Proc. ISSRE, 2012, pp. 271280.

[4] G. Antoniol, K. Ayari, M. Di Penta, F. Khomh, and Y.-G. Gueheneuc, Is it a bug or an enhancement?: a text-based approach to classify change requests, ACM, New York, NY, USA, , Article 23 , 15 pages. DOI: https://doi.org/10.1145/1463788.1463819

[5] John Anvik, Lyndon Hiew, and Gail C. Murphy. 2006. Who should fix this bug?. In "Proceedings of the 28th international conference on Software engineering " (ICSE '06). ACM, New York, NY, USA, 361-370. DOI: https://doi.org/10.1145/1134285.1134336

[6] K. Herzig, S. Just, and A. Zeller, Its not a bug, its a feature: how misclassification impacts bug prediction, in Proc. ICSE 13, 2013, pp. 392401.

[7] N. Pingclasai, H. Hata and K. Matsumoto, "Classifying Bug Reports to Bugs and Other Requests Using Topic Modeling," 2013 20th Asia-Pacific Software Engineering Conference (APSEC), Bangkok, 2013, pp. 13-18. doi: 10.1109/APSEC.2013.105