ISSN: 2586-7652 | 2635-7607 Vol. 02, No. 02 September 2019

International Journal of Advanced Engineering

Source: http://ictaes.org

Manuscript received: August 6, 2019; Revised: September 9, 2019 ; Accepted : September 20, 2019



Qualitative Technology Acceptance Evaluation of JIRA in Software Development Using Machine Learning

Ken Gorro[1], Kim Gorro[1], Anthony Ilano[1], Archival Sebial[1], Elmo Ranolo[1], Edita Vale[2]

[1]*University of San Carlos*
[2]*Abuyog Community College*

kdgorro, kidgorro, asilano, ajsebial, ebranolo{@usc.edu.ph}, valeedita@gmail.com

Abstract

In Software is an important thing that makes our life easier in day to day living. In creating a software in any field of expertise requires an engineering process and it includes software testing. Testing requires manpower that it needs bug monitoring tool. This study evaluates an existing bug tracking tool in qualitative approach. In this study, we ask 1000 software engineers as respondents to answer qualitative question. The respondents were asked the question why JIRA as bug tracking tool. This study uses qualitative approach in understanding the acceptability of a technology within organization. Word2Vec and Bi-term Topic Modeling (BTM) were used to analyze the responses. The Bi-Term topic models show the following important reason: project management, user experience, customizable and repository connectivity. BTM topics models shows 50.3% relationship between words using intrusion test. Word2Vec shows a significant result in project management and repository connectivity.

Keywords: *Bug tracking tool; JIRA;, Bi-term topic modeling; Word2Vec.*

1. Introduction

Software testing is an important step in software engineering process. The need for bug tracking tool in any software development project arises. The JIRA software was created by Atlassian [1]. Qualitative approach was used to gain further findings from the comments of the chosen sample. Analyzing each responses from various individual requires tremendous amount of effort. With the evolution of the technology, Natural

Corresponding Ken Gorro
Author's affiliation: 1University of San Carlos, Cebu, Philipine
Email: kdgorro, kidgorro@ usc.edu.ph

language processing techniques such as topic modeling and word2vec can analyze a given corpus to determine the latent topics and important key terms. Topic modeling is defined as a method for determining recurrent patterns of words in text. A number of topic modeling techniques and implementations exist. This study makes use of bi-term topic modeling and word2vec [2,3]. This study uses qualitative approach in determining technology acceptability over the traditional acceptance model. This study will be use to set a standard in Universities offering Software engineering program to include industry tools.

2. Related Study

Major advantages of BTM were highlighted in the study of yan et al. This study discusses the advantage of the LDA and PLSA [4]. Yan et al improve BTM and coined it Bursty Bi-term topic model (BBTM.) This was used in the twitter dataset and the result shows a significant improvement over the traditional topic models [5]. BTM was also used in headlines in comparison with LDA with full-text corpus and the result shows an equal with the LDA [6]. Another variation of BTM was introduced to surpass the limits of the traditional BTM and it was a Twitter-BTM model. The result shows a significant result that outperforms the traditional BTM [7]. Word2vec was utilized and improve in different research area in document ranking [8], sentiment analysis [9] and in named entity recognition [10]. Gorro et al, utilize word2vec and BTM in disaster risk suggestions to gain insights of the qualitative data [11].

3. Methodology

A.) Building corpus - Each response were treated as one document entry in the bi-term topic modeling.

B.) Data Cleaning - Irrelevant and noisy data is not helpful for text analysis. The following procedures were performed:

1.) Remove symbols

2.) Remove special characters

3.) Remove words with less than 4 characters.

C.) Bi-term topic modeling

The specific generative process of the corpus in BTM can be described as follows [4]:

1.) For each topic z:

- Draw a topic-specific word distribution $\varphi z \sim Dir(\beta)$

2.) Draw a topic distribution $\theta \sim Dir(\alpha)$ for the whole collection

3.) For each biterm b in the biterm set B

- draw a topic assignment $z \sim Multi(\theta)$

- draw two words: wi, wj $\sim$ Mulit(φz)..

To identify the best hyper parameters for BTM, series of experiments were conducted.

Table 1. Experiment data

Experiment Number	Number of topics (k)	Number of iterations	Alpha (50/k)	Beta
1	30	5,000	5	0.005

2	25	5,000	4	0.005
3	20	5,000	2.25	0.005
4	10	5,000	5	0.005
5	5	5,000	3.10	0.005

Based on the series of experiments conducted, experiment #5 has the best combination of hyper parameters in generating the topic models. The hyper-parameters are as follows: Alpha = 3.10 and Beta = 0.005. The number of (K) topics starts with 30 topic models and is decremented by 5 for each experiment until K = 5, totaling to 5 experiments conducted.

D. Open Coding

Open coding is a manual approach to analyze the context of qualitative data. In this study, open coding was used to gain further insights of the result of BTM and Word2Vec.

E. Word2Vec

For word2vec, this study uses gensim to generate the word embeddings. We manually examined the word which we think important and narratives was created per word embeddings.

4. Results and Analysis

Table 2. The topic models being generated from experiment #5

Topic labels	Topic models
Topic 0: JIRA in project management using EPIC feature.	Monitoring, project, management, epic, implementation, bug
Topic 1: JIRA features is customizable	Customize, workflow, accessible, watchers, monitoring, ratio
Topic 2: Training is necessary in using JIRA	Fields, status, training, needs, friendly
Topic 3: JIRA can be used together with another repository	Repository, connection, bitbucket, perforce
Topic 4: JIRA is accessible and report generation is easy	Accessible, searching, reports, status, monitoring

Based on the result, JIRA is not only a bug monitoring tool, it is commonly used now as project management tool. In the topic models, Customizable, repository connectivity and report generation are important features of JIRA for the respondents.

To evaluate the result of topic modeling, we perform intrusion test using human evaluators. A word is removed from the topic model and an intruder word, randomly selected from the words not part of the topic model, is added. This is repeated with different words being removed and added. This test evaluates the cohesion of the group of words belonging to the same topic in the topic model [11-13]. We use the following formula:

$$score = \frac{1}{n} \sum \frac{number\ of\ correctly\ identified\ word}{number\ of\ word\ groups}$$

$$where\ n\ is\ the\ number\ of\ human\ evaluators$$

In totality the intruder word that was identified correctly is 50.3%, this shows a strong relationship between words in the topic models.

To analyze further the comments from the respondents, we perform word2vec. Table 4 shows the word embeddings with the cosine similarity. We only choose words with high cosine similarity that is above 0.5..

Word	**Most similar words**	**Similarity**
Custom	JIRA	0.7092302
	Project	0.6030303
	Epic	0.7110340
	Implementation	0.5930232
Project	Epic	0.6130402
	Management	0.8032030
	JIRA	0.7210340
	Watchers	0.5694321
	Reports	0.6310392
Accessible	home	0.5139403
	internet	0.6128392
	monitoring	0.7312839
Friendly	user	0.8302939
	UI	0.7384021
	interface	0.6313422
	repository	0.9323824
	Perforce	0.6302193
	bitbucket	0.5943813

The following narratives were formed based on the word2vec result:

I.Project Management (Epic, Implementation, Reports, Monitoring)

The respondents give emphasis on the aspect of JIRA as project management tool and not only a bug monitoring tool. The respondents finds it useful the Epic feature of JIRA. Sample actual responses:

a) I like JIRA epic feature. It helps me prioritize critical ticket.

b) JIRA is not only for bug monitoring, we used in project management because it could be use for implementation ticket.

II. Repository Connectivity (bitbucket, perforce, repository

The respondents also give emphasis on the connectivity of JIRA on repository system which helps them monitor fixes. Sample actual responses:

a.) JIRA can be easily link to bitbucket.

b.) JIRA can customize and use it in perforce.

III. User experience (friendly, UI, user)

The respondents also give emphasis on the use experience of JIRA which helps them become more productive.

Sample actual responses:

a.) JIRA UI and user friendliness is amazing.

b.) Simple to use and it follow web standards.

In this study, analogy test was used to determine the relationship of BTM and Word2Vec result. We pick-up words from the topic models and perform analogy test with the word2vec result. The following relationships were found:

a.) repository + workflow - customizable = management

This gives us the insight that repository connectivity and workflow feature of JIRA are important features that should be customizable. Repository connectivity and workflow that can be customized are important to project management.

b.) User + Friendly - epic = project

This gives us the insight that user friendly is an important factor for users but without the epic feature of JIRA, the project management would still be at risk.

4. Conclusion

In this paper, we examined JIRA as an important tool in software development. Based on the result, JIRA is now widely used not only as a bug tracking tool but as a project management tool for software engineers. User friendly, User experience, Project management tool and Repository connectivity by JIRA are important factors that were also discovered using BTM and Word2Vec. Qualitative responses shows important points that helps determine the acceptability of a technology in an organization. This study gives the emphasis on the importance of qualitative data in understanding the acceptability of a technology within organization. This paper do away with the traditional acceptance model.

References

[1] Retrieve from: https://www.atlassian.com/software/jira

[2] Ding, C., & He, X. (2004). K-means Clustering via Principal Component Analysis. ICML '04 Proceedings of the twenty-first international conference on Machine learning, 29.

[3] Vector Representations of Words. (2017, August 17). Retrieved from TensorFlow: https://www.tensorflow.org/tutorials/word2vec

[4] Yan, X., Guo, J., Lan, Y., & Cheng, X. (2013). A Biterm Topic Model for Short Texts. Proceedings of the 22nd

[5] international conference on World Wide Web, 1445-1456.

[6] Yan, X., Guo, J., Lan, Y., Xu, J., & Cheng, X. (2015). A Probabilistic Model for Bursty Topic Discovery in Microblogs. Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, 353-359.

[7] Xia, Y., Tang, N., Hussain, A., & Cambria, E. (2015). Discriminative Bi-Term Topic Model for Headline- Based Social News Clustering. Proceedings of the Twenty-Eighth International Florida Artificial Intelligence Research Society Conference, 311-316.

[8] Chen, W., Wang, J., Zhang, Y., Yan, H., & Li, X. (2015). User Based Aggregation for Biterm Topic Model. Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Short Papers), 489-494.

[9] Nalisnick, E., Mitra, B., Craswell, N., & Caruana, R. (2016). Improving Document Ranking with Dual Word Embeddings. WWW'16 Companion. Liu, H. (2017). Sentiment Analysis of Citations using Word2vec.

[10] Siencnik, S. (2015). Adapting word2vec to Named Entity Recognition. Nordic Conference of Computational Linguistics, 239-243.

[11] Qualitative data analysis of disaster risk reduction suggestions assisted by topic modeling and word2vec Retrieved from http://ieeexplore.ieee.org/document/8300601/

[12] Surian, D., Nguyen, D. Q., Kennedy, G., Johnson, M., Coiera, E., & Dunn, A. G. (2016). Characterizing Twitter discussions about HPV vaccines using topic modeling and community detection. Journal of medical Internet research, 18(8)

[13] Konishi, T., Ohwa, T., Fujita, S., Ikeda, K., & Hayashi, K. (2016, February). Extracting search query patterns via the pairwise coupled topic model. In Proceedings of the Ninth ACM International Conference on Web Search and Data Mining (pp. 655-664). ACM.