

## Analyze the Positive Comments in the Stack Exchange

Xie Guojun<sup>1</sup> and J. Yamini Devi<sup>2</sup>

<sup>1</sup>*Department of Computer Science and Technology, Henan Polytechnic University,  
Jiaozuo, Henan, China*

<sup>2</sup>*Department of Computer Science and Engineering  
KL University, Vaddeswaram, Guntur*

<sup>1</sup>*guojun108@sina.cn, <sup>2</sup>j\_yaminidevi@yahoo.co.in*

### Abstract

*Today stack exchange is the most developing company from all over the world. Since 2012 the stack Exchange Company has tried to assess the positiveness of the comments done by the various users. Overall 7500 comments were made by users over 4 years. With our advanced system, we find positive comments from all over the world and find the exact positive comments based on the positive word string search. With our positive word bank, we can find the positive comments from the 7500 comments. In this paper, the performance measure is done in a dataset published in stack exchange and finding the comments and how frequently repeated user comments are. This is the tool to find positive and negative comments.*

**Keywords:** stack change, software

### 1. Introduction

Stack Overflow is an inexorably vital part of the online assets depended on by a large number of programming engineers [1]–[3]. Like most online groups, newcomer recognitions and future cooperation can be vigorously affected by their first associations with the group, particularly on the off chance that they are antagonistic. Having the capacity to consequently group remark agreeableness would permit designer groups like Stack Overflow to preemptively caution clients that endeavor to post remarks which might be gotten as antagonistic. This could diminish the number of unpleasant remarks while expanding support and eventually the life span of the group.

To examine whether Stack Overflow was inviting or threatening to newcomers, the Stack Exchange association affected a crusade in the late spring of 2012 to concentrate on and advance "superbness" in the Stack Overflow people group. Stack Exchange tested 7,000 of the more than 15 million remarks in the Stack Overflow database. Utilizing Mechanical Turk, these remarks were physically appraised for "kind disposition". The Stack Exchange consider found that Stack Overflow was turning into a friendlier place, with the rate of benevolent remarks expanding after some time. Be that as it may, they additionally found that first-time publications will probably get negative comments.

In this position paper, we play out an examination of the remark dataset distributed by Stack Exchange, concentrating on the substance of those remarks and how as often as possible individual remarks are rehashed. This is an initial move towards building instruments that naturally recognize conceivably negative remarks in designer groups.

---

#### Article history:

Received (February 5, 2016), Review Result (April 10, 2016), Accepted (May 10, 2016)

## 2. Analysis and findings

We played out a basic event examination to distinguish the number of times each of the 3,561 remarks showed up in the August 2012 Stack Overflow information dump and the number of one-of-a-kind clients who remarked. Table I introduces the aftereffects of this examination alongside an "excellence" arrangement because of the Stack Exchange contemplate. Most of the most much of time rehashed remarks are named "amicable" with just a single "unpleasant" remark being exceedingly rehashed. Altogether, we found more than 20,000 events of the 3,561 remarks. In any case, just 212 remarks were rehashed more than once. If the Stack Exchange dataset is illustrative of Stack Overflow remarks, the lion's share of remarks is one of a kind, however, a little number is very rehashed.

Another outstanding finding was remarked that happened much of the time, yet were made by just a little number of clients. For instance, Table II records a choice from the 100 most of the time happening remarks that each started from a solitary client (remarks have been condensed). Table III demonstrates the main 10 most now and again happening "disagreeable" remarks. Prominently, the "homework" watchword is not just present in the most rehashed "disagreeable" remark, but on the other hand, is available in half of the main 10 "antagonistic" remarks. This obliged vocabulary introduces a basic heuristic that could be utilized to hail "hostile" remarks.

Table IV demonstrates the most now and again happening equivocally arranged remarks. In contrast with the most much of the time happening "benevolent" or "hostile" remarks, equivocal remarks tend to be rehashed far less frequently and seem to have a wealthier vocabulary.

## 3. Existing system

Nowadays we have so many online communities. In those communities, we can log in and we can share our questions with others. But we do not have good communities for software developers. In normal communities, we can ask about normal questions only. In these communities, there are no software topics.

### *Disadvantage:*

1. There is no room for developers.
2. We cannot get correct answers from communities.

## 4. Proposed system

On this site, the admin can add categories. And subcategories for software developers. Users can create topics for exiting categories and users can create questions. And they can reply to those questions. All software developers can see answers to questions.

### *Advantage:*

1. This is for only developers.
2. All other users can give a proper answer to our questions.

### **Limitations:**

As we analyzed just remarks that stayed in the Stack Overflow database in August 2012, there is a probability that to a great degree negative remarks caught in the Stack Exchange study may have been erased before this date. Since the Stack Exchange dataset contained

remarks inspected at particular times more than four years, it is conceivable that some now and again happening remarks were not caught in that dataset, and accordingly would not be considered in our examination. At long last, the remark groupings given in the Stack Exchange dataset were gathered utilizing Mechanical Turk: members were not unequivocally prepared in remark characterization and did not have admittance to the setting in which a remark showed up.

## 5. Conclusions & future work

Using data published by Stack Exchange that rated the “friendliness” of Stack Overflow comments, we analyzed how frequently those comments were repeated on Stack Overflow and the vocabulary used in different types of comments. We found that the most frequently occurring “unfriendly” comments use a constrained vocabulary and that some very frequently repeated comments are created by small numbers of users. This qualitative analysis adds further insights to the quantitative analysis conducted by Stack Exchange so far. Our findings suggest possible features that could be implemented on Stack Overflow and other developer communities that would warn users that are about to post a comment that might be received as negative. In future work, we will investigate the automatic classification of comments and posts.

## References

- [1] B. Dagenais and M. Robillard, “Creating and evolving developer documentation: understanding the decisions of open source contributors”, in Proceedings of the Eighteenth ACM SIGSOFT International Symposium on Foundations of Software Engineering, ACM, pp. 127-136, (2010).
- [2] S.M. Nasehi, J. Sillito, F. Maurer and C. Burns, “What makes a good code example?: A study of programming q&a in StackOverflow”, in Software Maintenance (ICSM), 2012 28th IEEE International Conference on, pp. 25-34, Sept. (2012).
- [3] C. Parnin and C. Treude, “Measuring API documentation on the web”, in Proceedings of the 2nd International Workshop on Web 2.0 for Software Engineering, ser. Web2SE '11. New York, NY, USA: ACM, pp. 25-30, (2011). [Online]. Available: <http://doi.acm.org/10.1145/1984701.1984706>

***This page is empty by intention.***