

Identification of Vulgar Comments on social Media using data mining

Jash Parekh

NBN Sinhagad School of Engg.
Pune, Maharashtra, India
vroy1816@gmail.com

Harshad Shewale

NBN Sinhagad School of Engg.
Pune, Maharashtra, India
harshadshewale8397@gmail.com

Aman Mahajan

NBN Sinhagad School of Engg.
Pune, Maharashtra, India
amanmahajan78@gmail.com

Abstract- Data Mining is the process of finding anomalies, pattern and co- relations within large data sets to predict outcomes. These outcomes can then be manipulated to get our desired results and then work accordingly. Social media has greatly enriched people's lives, allowing them to share or post their feelings through posting various comments or pictures. Some friends or people comments are such vulgar that usually the person sharing the post deletes the post. Our approach is to detect such vulgar comments and delete them immediately, as soon as they are posted by someone and even after deleting such comments the user on who's post the comment was posted, gets to know the name of the person and comment which was posted. We are using Quick Sort algorithm for sorting comment. Through this project we aspire to remove negative comments and thus keep posts clean.

Keywords- Vulgar Comments, Post etc.

I. INTRODUCTION

In Existing System if one persons post a status or photo, Usually Friends pass comments and even other people who aren't your friends also pass comments on that post, the unknown people sometime comments vulgar comments and those comments can be seen immediately by the people who are active on social media. So we are developing a system that will automatically delete such comments immediately as soon as they are posted. the person on who's post the comment is done will receive a mail with that comment and get to know the name of the person and the comment who commented something vulgar on this post.

II. LITERATURE SURVEY

- 1. Sarita Yardi, Daniel Romero, Grant Schoenebeck-** In this paper a survey was made on spam and legitimate user, who tweets more spammers or legitimate spammers, who has more followers spammers or the legitimate user.
- 2. Miss Rohini D.Warkar, Mr.I.R.Shaikh -** In this paper a NPL (Natural Language Processing) technique was used for analyzing and representing, even the repeated comments on posts were removed by extracting the data from document and remove ambiguity in result.
- 3. Chenwei Liu, Jiawei Wang, Kai Lei -** In-depth analysis was made on the comment posted in Big-V user they constructed a Self Extensible spam Dictionary from a new perspective on semantic feature of words to detect

spam comments more accurately and effectively. They successfully identified spam comments to relatively high Precision Rate of 87.9% and recall rate of 87.2%.

III. PROJECT SCOPE

Now a day's Face book is a popular social media, where everyone is usually having an account. But some people feel unsafe to post a picture or status or to use face book because other people keep commenting on posts which can be vulgar. So we have developed a system through which such type of comments won't be shown and also get deleted by the system automatically and people can feel free and safe to post.

IV. SYSTEM ARCHITECTURE

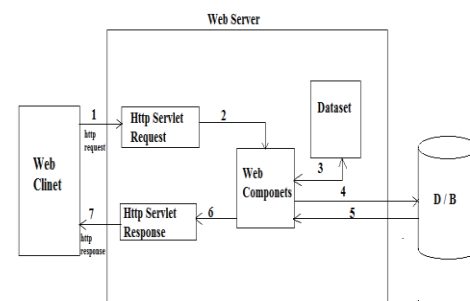


Fig.1 System Architecture.

V.CONCLUSION

We are conducted an in-depth analysis on the comments posted on social sites. To detect vulgar type comments more accurately and effectively, we are going to construct a Dataset / Dictionary and we are compare that dataset with the comments which can be a sentence or word and according to that in case we find any vulgar word in that sentence we are going to delete such kind of comments automatically as soon as they are posted. Hence we are removing all negative comments and thus all the post will be clean and positive.

ACKNOWLEDGMENTS

The volume of the work would not have been possible without contribution in one form or the other by few names to mention. We welcome this opportunity to express our heartfelt gratitude and regards to our project guide Prof. M.P. Navle Department of Computer Engineering, NBN SINHGAD SCHOOL OF ENGINEERING, PUNE-41 , for his unconditional guidance. She always bestowed parental care upon us and evinced keen interest in solving our problems.

REFERENCES

- [1]. Chenwei Liu, Jiawei Wang, Kai Lei :- Detecting Spam Comments Posted in Micro-Blogs Using the Self-Extensible Spam Dictionary , IEEE ICC 2016 ,SAC Social Networking . IEEE , 2016.
- [2]. Cristina Radulescu, Mihaela Dinsoreanu, and Rodica Potolea. Identification of spam comments using natural language processing techniques. In Intelligent Computer Communication and Processing (ICCP), 2014 IEEE International Conference on, pages 29–35. IEEE, 2014.
- [3]. Kate Raynes-Goldie. Privacy in the Age of Facebook: Discourse, Architecture, Consequences. PhD thesis, Curtin University, 2012.