

Abused Word Detection on Social Media

Tejaswini Yenurkar, Ankit Nandurkar, Nisha Thute, Prof. Rushikesh shete

Department of Information Technology

Datta Meghe Institute of Engineering, Technology and Research, Sawangi(Meghe), Wardha

Rashtrasant Tukdoji Maharaj Nagpur University

Abstract-Online social media now a days it's a medium to communicate each other and it's a platform of advertising, popularity and so on. It offers a platform for the people to connect online and also gives the privacy for one-to-one interaction. The different type of languages having different type of abused words that's having different typos is really hard to recognise the word that's the big problem. Most of the people using abused words on social media. By this the social environment become polluted or unhealthy for young people. It gives the bad impact on students and their mind. For that we proposed a system which is hiding the abused word on the social media without exposing publically to decrease the death rate of students which is harassed on social media and commenting negative comments social media to overcome this issue.

I. Introduction

Social Media, as defined in is a group of Internet based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user-generated content. Via social media, people can enjoy with different type of information, convenient communication experience and so on. However, social media may have some side effects such as cyberbullying, which may have negative impacts on the life of people, especially children and teenagers.

Cyberbullying can be defined as aggressive, intentional actions performed by an individual or a group of people via digital communication methods such as sending messages and posting comments against a victim. Different from traditional bullying that usually occurs at school during face-to-face communication, cyberbullying on social media can take place anywhere at any time. For bullies, they are free to hurt their peers feelings because they do not need to face someone and can hide behind the Internet. For victims, they are easily exposed to harassment since all of us, especially youth, are constantly connected to Internet or social media.

By the research on the social media it gives the result of analysis which is said that the users on the social media chat with friends personally and also within the group which formed by the other users and in that the users are using the abused words in the conversations between them by this the social environment become unhealthy and gives bad impact on the youngsters. By the analysis 43% teenagers are bullied on social media by this teenagers are attempting self-injuries and suicide message. For efficient process they concentrate on the text content. It is more reliable term. This paper focuses on the social media user's latency of negative comments on the public post area, social environment, the abused word doesn't get exposed publically and also focuses on the adult image which posted on the social media doesn't exposed publically.

II. Literature Review

Rui Zhao and Kezhi Mao” Cyberbullying Detection based on Semantic-Enhanced Marginalized Denoising Auto-Encoder”

This paper addresses the text-based cyberbullying detection problem, where robust and discriminative representations of messages are critical for an effective detection system. By designing semantic dropout noise and enforcing sparsity, they have been developed semantic-enhanced marginalized denoising auto encoder as a specialized representation learning model for cyberbullying detection.

Shambhavi Dinakar, Pankaj Andhale, and Manjeet Rege Rochester Institute of Technology, Rochester, NY Intuit, Inc., Mountain View, CA University of St. Thomas, St. Paul, MN “SentimentAnalysisofSocialNetworkContent” 2015 IEEE 16th International Conference on Information Reuse and Integration”.

This paper addresses the semantic analysis of the comments which is on the social media that shows the negative comments and also reported that comments ratio in social public area. Due to sparsity constraints the lack of bullying data is present, so the analysis hard to described the proportion of the cyberbullying. In this analyse the psychological behaviour of the victims.

Summary

Both paper are related to cyberbullying by their algorithm of detecting abused words and another is to analyse the semantic words, and also analyse the psychological behaviour. By this we referred that information and analysis we implement a web application. The referred proposed algorithm it detects abused word automatically, but it takes more time to detect exact word by their matching percentage. This issue we implement an

algorithm to increase the speed of detecting the abused word.

III. Methodology

A. Data Source

We studied the abuse words which is used by the people, it has millions of users there are increasing number of abused words are used by the various users publically and adult images. The social websites are having many abused words which used by peoples, we observing the tweets or comments and inserting that abused words in our own created data set.

B. APRIORI Algorithm

Apriori is an algorithm for frequent item set mining and association rule learning over transactional databases. It proceeds by identifying the frequent individual items in the database and extending them to larger and larger item sets as long as those item sets appear sufficiently often in the database. The frequent item sets determined by Apriori can be used to determine association rules which highlight general trends in the database. In this paper proposed to extract the frequent item sets in the larger abuse words. It uses breadth-first search and a Hash tree structure to count candidate item sets efficiently.

Pseudo-code:

C_k : Candidate itemset of size k

L_k : Frequent itemset of size k

$L_1 = \{\text{Frequent items}\};$

for($k=1; L_k \neq \emptyset; k++$) do begin

$C_{k+1} = \text{Candidates generated from } L_k;$

 for each transaction t in database do

 Increment the count of all candidates in

C_{k+1}

 that are contained in t

$L_{k+1} = \text{Candidates in } C_{k+1} \text{ with min_support}$

 end

return $U_k L_k$;

Above shows the Pseudo-code of an algorithm which is used to determine the words are mostly used in the socially. To rapid recognition it helps to give robustness in the finding the words in database Apriori uses. The pseudo code for the algorithm is given below for a transaction database $\{T\}$, and a support threshold of $\{\epsilon\}$. Usual set theoretic notation is employed, though note that $\{T\}$ is a multiset. C_k is the candidate set for level $\{k\}$. At each step, the algorithm is assumed to generate the candidate sets from the large item sets of the preceding level, heeding the downward closure lemma. $\text{count}[c]$ accesses

a field of the data structure that represents candidate set $\{c\}$, which is initially assumed to be zero. Many details are omitted below, usually the most important part of the implementation is the data structure used for storing the candidate sets, and counting their frequencies.

C. Bag of words

We use the concept, this are taking the whole datasets directly which readymade data set on the web. The bag-of-words model is a simplifying representation used in natural language processing and information retrieval (IR). Also known as vector space model. In this model, a text (such as a sentence or a document) is represented as the bag (multiset) of its words, disregarding grammar and even word order but keeping multiplicity. The bag-of-words model has also been used for computer vision. The bag-of-words model is commonly used in methods of document classification where the (frequency of) occurrence of each is used as a feature for training a classifier.

D. Data Aggregation

We use to boosting extraction the whole comment into number of parts means to divide the words by their initial letter to inserting into the specific cluster to separate the word too easily and fast to identify the abused word. The every comment would through that process after it goes on the comment area. Data aggregation is the compiling of information from databases with intent to prepare combined datasets for data processing. The source information for data aggregation may originate from public records and criminal databases.

Data Analysis

The goal of data analysis described is to examine the abused comments which contains the abused words, we analysed the abused words and finding which type of words and types are they used. We analysed the comments which is posted by the user on our web application. We acquired result as the user using short forms of words that's why it's quite difficult to trace by our algorithm. But it not possible to recognise automatically. The number of users are using abused words in the larger amount than expected, that inspecting the amount of comments with included their short forms of abused words. By this we get conclusion is that the abused words are not absolute and not limited they are infinite typos of single word.

Working

A. Web Application

We implemented such social web application like "twitter" and named as "My Tweet" it having the public comment area and sub twitting feature and so on. In that we providing the comment feature, report about the user, suggest about the new abused words in the comments, uploading images

on wall and deleting comment which commented by itself. In our web application we try to providing a healthy social environment by hiding abused words without exposing publically.

In this web application user use the abused words in the comment, he can comment also but that abused word doesn't shown publically, it will be hide from everyone only that word seen by which is commented abused word. After that if any user irritate by this type of commenting it can be report about that user to admin, then admin decides block the user or not. In that any user can suggest the abused word which is not having in our data set. The reported user firstly examined by the admin and after that it will take action on them.

The suggestive word is added by the admin in our dataset, the suggestive word should represented by "-word-" this the abused word should be placed between the dashes. By this it is to identify the admin any abused word suggest by the user. And lastly about images the user can upload any image but firstly it goes to accept that image it will be shows or appear on the public wall, in other case the if image rejected that doesn't appear on the public wall.

B. Word Matching

We used such algorithm to find an exact abused word among the number of words, we simply comparing the word with the words word's and match them. If it is matched it simply replaced by the special character i.e. "*****" that star replaced by the whole word which is matched our dataset words. If the number of words are existing the single comment in that time the data aggregation concept is help to find abused word among them faster and accurate. Every comment to post firstly it goes through our algorithm and after that it will be appear on the public wall. If any comment contains any abused word it simply recognised that word and replaced by the stars by this any abused word doesn't exposed publically.

IV. Conclusion

This paper addresses such word matching algorithm to find faster and accurate abused words the and how to hide the abused words on the public posting area and provide the healthy social environment on our social web application. In the data analysis we examined the every user's comments to survey on the abused words to increase the number of abusing words in our dataset. By this we identifying the aggressiveness of user and the harassment level of people by social media. The different types words are not detected by our algorithm it's a tough job to detect. And that's the futuristic scope of that to identify the exact word in the whole sentence.

References

- [1]. M. Kaplan and M. Haenlein, "Users of the world, unite! The challenges and opportunities of social media," *Business horizons*, vol. 53, no. 1, pp. 59–68, 2010.
- [2]. R. M. Kowalski, G. W. Giumetti, A. N. Schroeder, and M. R. Lattanner, "Bullying in the digital age: A critical review and meta-analysis of cyberbullying research among youth." 2014.
- [3]. Shambhavi Dinakar*, Pankaj Andhale†, and Manjeet Rege‡ *Rochester Institute of Technology, Rochester, NY †Intuit, Inc., Mountain View, CA ‡University of St. Thomas, St. Paul, MN "SentimentAnalysisofSocialNetworkContent" 2015 IEEE 16th International Conference on Information Reuse and Integration.
- [4]. Rui Zhao and Kezhi Mao" Cyberbullying Detection based on Semantic-Enhanced Marginalized Denoising Auto-Encode.