International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 8 (2015) pp. 18911-18916

http://www.ripublication.com

Analysis of Twitter Feeds Using Natural Language Processing And Machine Learning

Harish A. Raman, MuraliKrishnan.E, Abishek.M, SaiSandhesh.R, Karthik Vijaykanth, Harini.N
Dept. of Computer Science and Engineering
Amrita School of Engineering, Coimbatore
Amrita VishwaVidyapeetham (University),India

Abstract

Cyber bullying is a rapidly burgeoning phenomenon in to-days world dominated by the Internet. From every major incident happening around the world to meager day-to-day activities of an individual is posted on social media. Ergo, Internet has now become an essentiality that is indispensable. Though this seems intriguing, however, it has led to the advent of cyber bullying. Social networking sites provide an easy platform for the cyber bullies to identify and victimize other users. Cyber bullies may make use of victims personal data(e.g. real name, home address) to impersonate them, or by creating fake accounts in social networking sites that defames, discredits or ridicules them. Due to the anonymity of the Cyber bullies it becomes increasingly difficult for the o ender to be caught and punished for their behavior. This paper proposes a system which identifies posts which are aimed at hurting the sentiments of other users and makes the user to rethink and hence refrain from posting the same. This paper also provides an effective algorithm that identifies and reduces the spam content in the users post/tweet.

Keywords: Cyber Bullying,post/tweet, NLP

Introduction

In the era of Internet technologies, the field of social media and analysis have gained prominence.93% of young people between the ages of 12 and 17 were online. In fact, youth spend more time with media than any other single activity besides sleeping.[1] It also suggests that a great deal of people follow celebrities and track celebrities in such Social Networking sites. An important threat for people with many followers is that some of their posts contain sensitive informa-tion which may harm many. To avoid the circumstance of people getting hurt, such posts are identified and removed through a polling. Such crowd-sourced so-lutions may not end up with a scenario

where a harmful post is removed. To avoid such scenarios where human interaction is required, a machine learning solution involving natural language processing is adopted. In scenarios mentioned above, there is a possibility that communication through Twitter or Facebook is used to bully another person. Survey shows that more than a million children were harassed, threatened or subjected to various forms of cyber bullying on social networking sites such as Facebook and Twitter during the last year, while 90% of teenagers who use social networking sites who have witnessed online cruelty say they have ignored mean behavior on social media, and the 35% have done this frequently. Also, 55% of the 95% of teenagers who use social networking sites have frequently witnessed others ignoring mean behavior in these sites.[2] This is called cyber bullying and has grown to become a global problem. Currently, there is no existing solution to prevent cyber bullying from occurring. Features like mute, block and report abuse do not negate the act of cyber bullying from happening, they are only helpful in preventing further cyber bullying from occur-ring temporarily as it is only a matter of time before the potential cyber bullies start showing up in other social networking sites against same/potential victims making the solution ineffective in the long run. This paper proposes a system which acknowledges the above shortcomings and provides an efficient solution by identifying potentially harmful posts before actually posting it. It also provides a method to identify posts/tweets that advertises a product thereby effectively reducing spam content in Social Networking sites. The main contributions of this work include:

1. Creating an extension for Google Chrome that acts as a gateway for access of tweets/posts.
2. Detecting harmful messages using sentiment analysis through NLP and Ma-chine Learning techniques.
3. Adding features for identifying and removal of tweets involving buying and selling of a product.

Related Work

With the ever increasing popularity of Social Networking sites[3], sentiment anal-ysis of Twitter/Facebook feeds are starting to become a field of interest in the research literature. The research ranges from document level classification[4] to finding the polarity of words and phrases of a given sentence. Taking into account of the word limit of 140 characters in Twitter, sentiment analysis of Twitter feeds are almost similar to that of sentence level sentiment analysis.[6]

For subjective data the usual method is to collect twitter data that end with emoticons such as Go et al., (2009). They treat tweets that end with positive emoticons like :-) as positive, and negative emoticons like :-(as negative. Kouloumpis et al., 2011 Davidov et al., 2010; Agarwal et al., 2011 make use of n-grams and polarity lexicons in determining the sentiment of a given tweet/-post. Nevertheless, in all the given approaches, the contextual information of a tweet/post is not exploited but rather the lexicons features are used. For objective data, popular newspaper accounts in Twitter such as The Hindu or Times of India are crawled and n-gram

models are used to determine the sentiment. Since the data that is obtained by crawling the Twitter feeds with the help of search queries, they are mostly biased.

Lexicon-based methods use the orientation of words found in a given sentence to calculate the overall sentiment. The lwall et al. proposed Senti Strength; a lexicon-based method to detect the sentiment on Social Networking sites. This method resolves the ill-formed language usually present in social networking sites such as Twitter by applying several lexicon rules, and identifying the presence of emoticons, negations and booster words. The main drawback of this lexicon-based methods is that the full dependence on the presence of words or syntactic features that are mainly used to predict the sentiment of the given sentence. In many cases, however, semantics of the context plays a rather important role than the individual presence of a word. This paper acknowledges the same, and ensures both the context and the individual presences of the words in a sentence are taken into account. Also, many papers usually classify a given sentence into three different categories viz. Positive, Negative and Neutral. In contrast, we take a step further and try to predict the probability of the usage of the Neutral sentence in that context as to whether it is more positively oriented or negatively oriented.

System Architecture

In this section we present the architecture of the proposed system for identifying and classifying the tweets into various categories and filtering those contents that involve buying and selling of a product before they are posted as tweets. The proposed system consists of four major processes which can be identified as:

1. Retrieving tweets from the browser: First, the tweet that is to be posted by the user, is retrieved dynamically from the tweet-box/tweet-area with the help of an extension designed for Chrome browser and is sent to the server for analysis.

Figure 1: Screenshot

2. Classifying the tweets into various categories: Once the tweet that is to be analyzed is retrieved, it undergoes a sentiment analysis using Stanford CoreNLP package which makes use of various annotations such as lemmatization, POS tagger and then adds the corresponding sentiment of the given sentence. The package classifies the sentiment of the given sentence into three different categories viz. Positive, Negative and Neutral. Once the sentiment is known, corresponding alerts are sent back to the user for positively and negatively classified sentences.

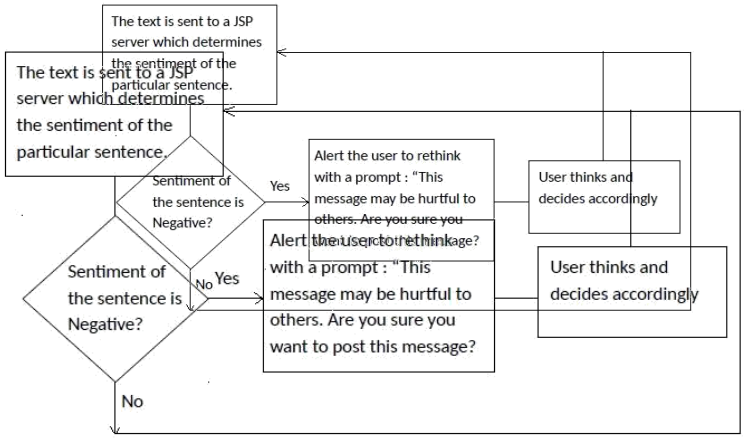

Figure 2: Flow Chart

3. Reclassification of Neutral sentiment tweets: In the case of sentences that are classified as neutral, it undergoes further analysis where the features of the given sentence is extracted and Nave-Bayes classification is performed to find the probability that the given sentence is negative.
4. Spam Detection: The algorithm retrieves the tweet from the user as input and notifies back to him whether it is spam or not by identifying those tweets that contain URL to any ecommerce or Online shopping website. It is assumed that any tweet/post that contains such URL involves buying or selling of a product. This is achieved with the help of SimilarWeb API that categorizes an URL into various categories and sites that fall under the category of ecommerce or Online shopping are identified, and an appropriate alert is sent back to the user notifying the same.

Results

A dataset of 120 tweets were taken, and were analyzed using the Stanford Core NLP package. The format of the dataset was really simple consisting of only the message written in the tweet box in English language. Additional information such as time when the tweet was written and the username of the author of the tweet were not taken into consideration. Here, HC denotes Human computed values done subjectively, and MC denotes the Machine computed values.

The result of the analysis yielded the following:

Table 1:

Number of tweets	HC Positive	HC Negative	HC Neutral
MC Positive	47	13	6
MC Negative	6	28	1
MC Neutral	2	11	6

The system observed, shows an accuracy of 81%.In contrast, the average accuracy of the Senti WordNet average scoring approach used in Sentiment Polarity Analysis was 51%.[10] We analyze from the above table that 11% of the tweets computed subjectively are mapped to Neutral, in comparison to Machine computed 16%. Out of this 16%, 9% of the tweets were subjectively mapped as Negative. This was a result due to random sampling. Thus, there is a need to ac-count for this 9% for a more efficient computation, thereby aiding in identifying tweets that may aid Cyber bullying.

Conclusion and Future Work

Our study involved analyzing tweets/posts to identify hazardous tweets that may aid Cyber bullying. We show that approximately 45% of posts on the Internet are negative or mixed in nature. Out of this, through our system we can remove about 35% of the tweets, i.e those that are negative and those that are slightly negative in nature. The proposed system aims to achieve a higher accuracy than obtained by Stanford CoreNLP, and aid in identifying tweets/posts that are neutral and tend towards negative.

Once these sentences that are classified as Neutral by the classifier are col-lected as a corpus, a machine is designed and trained so as to effectively predict and further classify the given neutral sentence. Features of the sentence are then extracted, and N-grams are constructed out of them. Unigrams and Bigrams of the sentence is then used to analyze and predict the probability whether the given neutral sentence is positively oriented or not. Polynomial Nave-Bayes classifier is used to build the second phase of sentiment classifier. Nave-Bayes classifier is based on the Bayes theorem (Anthony J, 2007).

$$P(M) = P(s)* (P(s)/P(M))$$

where s is a sentiment, M is the given neutral sentence.

Further, Twitter and Facebook are also used as a medium to advertise various products. Analyzing such tweets prove to be an additional task, thus reducing the performance of the system. The proposed Spam detection method provides a way to recognize such tweets and help in reducing the spam content in Social Networking Sites.

References

[1] Lenhart, Amanda, Kristen Purcell, Aaron Smith, and Kathryn Zickuhr. "Social Media & Mobile Internet Use among Teens and Young Adults. Millennials." Pew Internet & American Life Project (2010).

[2] Lenhart, Amanda. "Teens, smartphones & texting." Pew Internet & American Life Project (2012).

[3] Leading social networking sites. http://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/
[4] Pang, Bo, and Lillian Lee. "Opinion mining and sentiment analysis." Foundations and trends in information retrieval 2, no. 1-2 (2008): 1-135.
[5] Go, Alec, Richa Bhayani, and Lei Huang. "Twitter sentiment classification using distant supervision." CS224N Project Report, Stanford (2009): 1-12.
[6] Agarwal, Apoorv, BoyiXie, Ilia Vovsha, Owen Rambow, and Rebecca Passonneau. "Sentiment analysis of twitter data." In Proceedings of the Workshop on Languages in Social Media, pp. 30-38. Association for Computational Linguistics, 2011.
[7] Kouloumpis, Efthymios, Theresa Wilson, and Johanna Moore. "Twitter sentiment analysis: The good the bad and the omg!." ICWSM 11 (2011): 538-541.
[8] Davidov, Dmitry, Oren Tsur, and Ari Rappoport. "Enhanced sentiment learning using twitter hashtags and smileys." In Proceedings of the 23rd International Con-ference on Computational Linguistics: Posters, pp. 241-249. Association for Compu-tational Linguistics, 2010.
[9] Hayter, Anthony. Probability and statistics for engineers and scientists. Cengage Learning, 2012.
[10] Denecke, Kerstin. "Using sentiwordnet for multilingual sentiment analysis." In Data Engineering Workshop, 2008. ICDEW 2008. IEEE 24th International Conference on, pp. 507-512. IEEE, 2008.
[11] Lin, Chenghua, and Yulan He. "Joint sentiment/topic model for sentiment anal-ysis." In Proceedings of the 18th ACM conference on Information and knowledge management, pp. 375-384. ACM, 2009.
[12] Bermingham, Adam, and Alan F. Smeaton. "Classifying sentiment in microblogs: is brevity an advantage?." In Proceedings of the 19th ACM international conference on Information and knowledge management, pp. 1833-1836. ACM, 2010.