www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 7 Issue 1 January 2018, Page No. 23386-23489
Index Copernicus Value (2015): 58.10, 76.25 (2016) DOI: 10.18535/ijecs/v7i1.03

Review On: Detection of Spam Comments Using NLP Algorithm

Miss Rohini D.Warkar[1], Mr.I.R.Shaikh[2]

[1,2] Computer Engineering Department, Savitribai Phule University,Pune,India,

Abstract:
Detecting trending topics is perfect to summarize information getting from social media. To extract what topic is becoming hot on online media is one of the challenges. As we considering social media so social services are opportunity for spamming which greatly affect on value of real time search. Therefore the next task is to control spamming from social networking sites. For completing these challenges different concepts of data mining will be used. For now whatever work has been done is narrated below like spam control using natural language processing for preprocessing and clustering. One account has been created for making it real.

Keywords: Control Spamming ,Text mining ,Information filtering, Social Networking site, NLP .

1. Introduction

Now a days there is popularity and importance of social media sites are enhanced in people's daily life. Social media allows online users to share their feelings through posting comments. However, more and more spam comments are also being posted in user's account on the social media. So necessity of spam detection is increased. In traditional system, there are different systems that are used for spam detection such as The Naive Bayes classifier and tfidf(term frequency – inverse document frequency) . But these methods do not take the semantic information of the spam words or phrases into account, which leads to incomplete results.
Therefore, the requirement of research is to take into full account the significance of the semantic information of the words within all comments posted, including the vector expression of a word. And the vector distance between words. The skill of mining additional semantic features from words has been widely used in emotional classification and text classification, both of which have achieved good results.

This problem is solved by detecting the spam comments posted on social media site, through a combination of methods based on a deep learning model and statistical analysis. The Self-Extensible Spam Dictionary employs the deep learning Model of Skip-Gram, whose process of building is divided into three progressive stages:

(1) Subjectivity Judgment:-It is use to find out the semantic distinctions of words, dividing each word into either normal or spam;
(2) Category Judgment:-It is used to demarcate a word or phrase from the comments as an AD or vulgar category;
(3) Weight Judgment:-It is used to measure the extent of subjectivity and category, that is, the spam extent of a word or phrase in the AD or vulgar category.

The Proportion-Weight Filter Model uses statistical analysis to select the proportion and total weight of spam words contained in a single comment as the two key factors in deciding whether the comment is spam or not.
This model addresses the problem that the distribution of spam words in short and long comments is different. If we only detect spam comments by the factor of spam-weight, the longer of the normal comment, the greater the likelihood of matching a more low-weighted word in the spam dictionary is. This can cause the spam-weight of the normal comment to be high, which can then in turn reduce the Precision Rate.

Moreover, the shorter length spam comments with just one or two middle-weighted spam words will cause the total weight to be lower than the standard line, and thus identify the comment as normal, and in turn reduce the Recall Rate. Through a combination of the two critical factors, we can obtain more accurate results for detecting spam comments.

The improvement in the system is acquired by using Natural language processing (NLP) technique.

NLP Technique:

NLP is the technique that belongs to the CS taxonomy as the child of Artificial Intelligence (AI). Natural Language Processing is a technique used for analyzing and representing naturally occurring texts at one or more levels of linguistic analysis for the purpose of achieving human-like language processing for a range of tasks or applications. "Naturally occurring texts" can be of any language, mode, genre, etc. The texts can be oral or written and must be in a language used by humans to communicate to one another.
Significantly the text being analyses should not be specifically constructed for the purpose of the analysis, but rather it should be collected from actual usage. In simple terms, NLP is the use of computers to process written and spoken language for some useful purpose: to translate languages, to get information from the web on text data banks so as to answer questions, to carry on conversations with machines. Natural language processing approaches fall roughly into four categories: symbolic, statistical, connectionist, and hybrid.

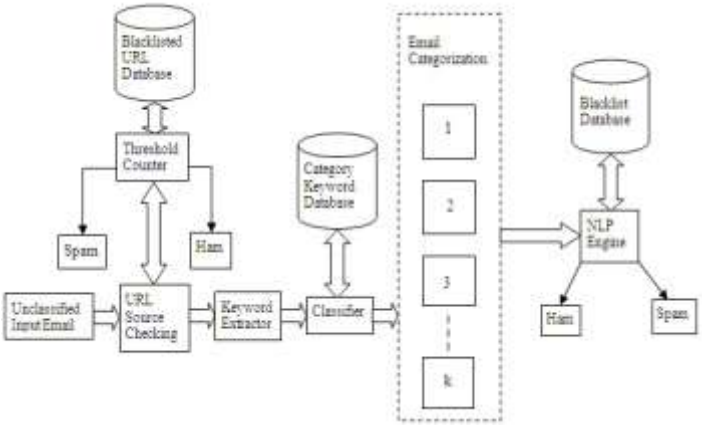

Figure 1: Design of NLP Processing

2. Literature Review

1.M. Cataldi, C. Schifanella and L. Di Caro proposed two measures, term frequency to calculate nutrition for each word and a page rank measure. After that Bursty keywords are obtained using nutrition trend. Then by using graph based approach for bursty keywords generates the topic boundary. analyze how the numeric features such as new lines, punctuation marks, links, white spaces, capital letters, vulgar words help eliminate incoherent and offensive comments and how the topic of the comment influences the detection of spam comments.

2.Sayyadi, Maykov and Hurst used graph approach in which clustering of keywords is done by matching pairs. They used community detection algorithm in which made a graph whose nodes are clustered. Also the topic extraction is carried out by identifying document with similar term. In this paper we presented a novel approach to detect in real-time emerging topics on Twitter. We formalized the keyword life cycle leveraging a novel aging theory intended to mine terms that frequently occur in the specified time interval and they are relatively rare in the past. We also studied the social relationships in the user network in order to quantify the importance of each analyzed content.

3.Lehmann, Kleinberg and Backstorm have used the graph for short phrases. Phrases are connected by edges developed a framework for tracking short, distinctive phrases that travel relatively intact through on-line text and presented scalable algorithms for identifying and clustering textual variants of such phrases that scale to a collection of 90 million articles, which makes the present study one of the largest analyses of on-line news in terms of data scale.Our work offers some of the first quantitative analyses of the global news cycle and the dynamics of information propagation between mainstream and social media.

4. David M. Blei,Andrew Y. Ng,Michael I. Jordan[5] used Latent Dirichlet allocation (LDA) is a topic model that generates topics based on word frequency from a set of documents. LDA is particularly useful for finding reasonably accurate mixtures of topics within a given document set.

3. Existing System

In existing system, structured methods of identifying spam comments, such as the Naive Bayes classifier, and tfidf (term frequency - inverse document frequency).It do not take the semantic information of the spam words or phrases into account, which leads to an incomplete results. Therefore, our research attempts to take into full account the significance of the semantic information of the words within all comments posted, including the vector expression of a wordand the vector distance between words. The skill of mining additional semantic features from words has been widely used in emotional classification and text classification, both of which have achieved good results .

Disadvantages:

1. Not able to detect non English Words and spam messages which are encrypted
2. Incomplete spam selection as semantic analysis not considered.

4. Proposed System

There are two approaches in the proposed work, identifying current and control spamming. The first step is pre-processing which is important for mining the data or filtering the data. The work of pre-processing has been done. Then the spam control has been done. Spam control is the part of feature extraction. Here used the bisecting K-means clustering algorithm, because clustering is an important step for quality results. So nothing but natural language processing (NLP) technique has been used for pre-processing, clustering etc.

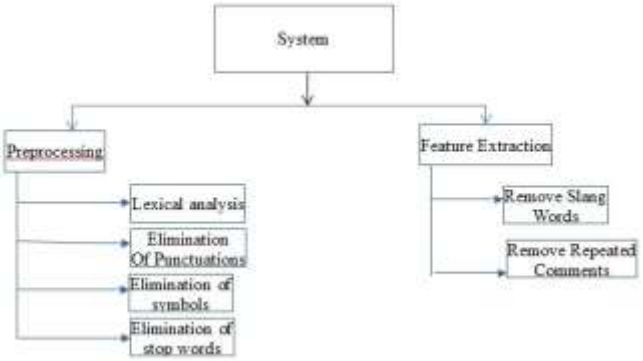

Figure 2:Design of system

1. Preprocessing
2. Feature Extraction

Preprocessing

Preprocessing contains filtering of data. Natural language processing concepts are used for preprocessing of data.

Lexical analysis

The lexical analyzer covert sentences into words then words convert into characters.

Elimination of punctuations

Remove punctuations like comma, full stop etc.

Elimination of symbols

Remove symbols like @, # etc.

Elimination of stop words

Remove words like in,of, the, is, and, for etc.

Feature Extraction

Feature extraction is used for reduction of dimensionality. Before classification there is need ofreduction of feature space.

Now spam control is also nothing but a feature reduction task. Therefore, slang word reduction is done for the spam control.

Remove of slang word

For spam control, dictionary of slang words is created. So, whenever user use any slang word in the post or comment that word matches with the words available in the dictionary and it replaces with the stars (****).

Remove Repeated comments

In extracting the data from document remove ambiguity in result.

5. Algorithm

The Natural Language Processing(NLP) is used to processing the data to find spam detection and also /iterative algorithm is used for constructing the spam dictionary.

NLP Algorithm:

It contain flow of different phases given below:

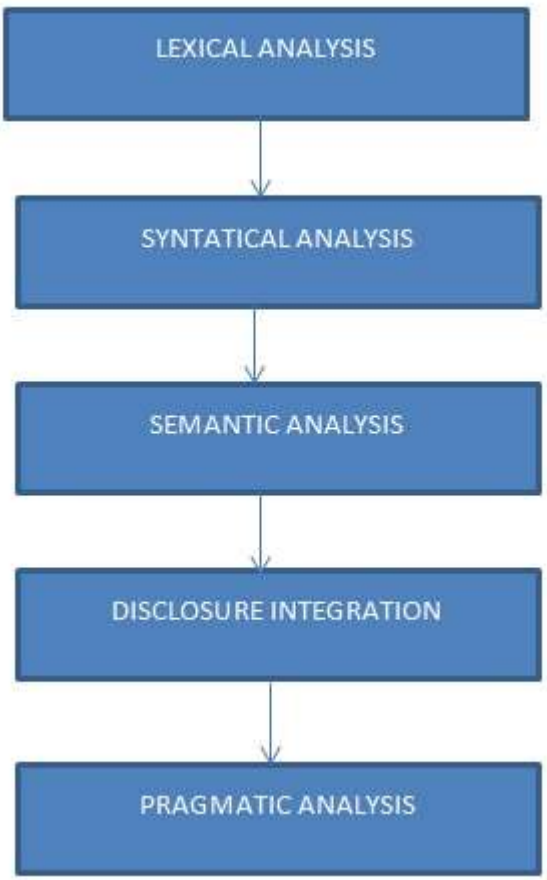

Figure 3: NLP Algorithm

Iterative Algorithm

Iterative algorithm for constructing the Domain Spam Dictionary

1. **Procedure** : Construct (Spam words)
2. **Input**: spam words of Basic vulgar and AD dictionary or added in the result of the previous iteration.
3. Acquire 15 most similar words for word s by comparing the semantic similarity between them;
4. And Add the word s_ into the candidate spam dictionary;
5. Delete words of candidate dictionary if they exist in Basic vulgar and AD dictionary or DS dictionary;
6. Calculate the avg-weight of same words in candidate spam dictionary.
7. for each word s__ in candidate spam dictionary do Acquire 15 most similar words for word s__ by comparing the semantic similarity between them;
8. if there are more than 4 words among exist in Basic vulgar and AD dictionary then Add word s__ into DS dictionary; else Drop it;
9. Empty the candidate dictionary
10. **Output:** the newly added spam words in this iteration.

Advantages:

1. As the system uses NLP it detect non English Words and spam messages which are encrypted.

2.The semantic analysis and Proportional weight filter technique do the complete spam selection.

6. Conclusion

The main aspects of the proposed work are to detect the current topics of real world and to control the spamming created by spammer. Pre-processing process is done. One account is created for showing results Also feature extraction is the part of spam control has done.

References

1. Chenwei Liu, Jiawei Wang, Kai Lei, "Detecting Spam Comments Posted in Micro-BlogsUsing the Self-Extensible Spam Dictionary", IEEE ICC 2016 SAC Social Networking
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. M. Cataldi, L. Di Caro and C. Schifanella, "Emerging topic detection on Twitter based on temporal and social terms evaluation," in Proc. MDMKDD: 10th Int. Workshop Multimedia Data Mining, New York, NY, USA, 2010, pp. 4:1–4:10, ACM.
4. Sayyadi, M. Hurst and A.Maykov, "Event detection and tracking in social streams," in ICWSM, E. Adar, M. Hurst, T. Finin, N. S. Glance, N. Nicolov, and B. L. Tseng, Eds. Palo Alto, CA, USA: AAAI Press, 2009.
5. J. Leskovec, L. Backstrom, and J. Kleinberg, "Meme-tracking and the dynamics of the news cycle," in Proc. KDD: 15th ACM Int. Conf. Knowledge Discovery and Data Mining, New York, NY, USA, 2009, pp. 497–506.
6. D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," J. Mach. Learn. Res., vol. 3, pp. 993–1022, Mar. 2003
7. Joseph Lilleberg, Yun Zhu, and Yanqing Zhang,"Support vector machines and word2vec for text classification with semantic features".,In Cognitive Informatics & Cognitive Computing (ICCI*

CC), IEEE 14th International Conference on, pages 136–140, 2015.
8. Bai Xue, Chen Fu, and Zhan Shaobin,"A study on sentiment computing and classification of sina weibo with word2vec." In Big Data (BigData Congress), IEEE International Congress on, pages 358–363 IEEE, 2014.
9. Huiyu Wang, Kai Lei, and Kuai Xu, "Profiling the followers of the most influential and verified users on sina weibo.", In Communications (ICC), IEEE International Conference on, pages 1158–1163. IEEE, 2015.
10. Ala' M. Al-Zoub, Ja'far Alqatawna, Hossam Faris,"Spam Profile Detection in Social Networks Based on Public Features", 8th International Conference on Information and Communication Systems (ICICS),2017.
11. Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon ," What is Twitter, a Social Network or a News Media?", In Proceedings of the 19th international conference on World wide web, pages 591–600. ACM,2010
12. Rohit Giyanani, Mukti Desai,"Spam Detection using Natural Language Processing",IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 16, Issue 5, Ver. IV (Sep – Oct. 2014), PP 116-119
13. Cristina Rădulescu , Mihaela Dinsoreanu , Rodica Potolea ,"Identification of Spam Comments using Natural Language Processing Techniques", IEEE 2014
14. Zengcai Su , Hua Xu;_ , Dongwen Zhang and Yunfeng Xu ," Chinese Sentiment Classification Using A Neural Network Tool - Word2vec", InMultisensor Fusion and Information Integration for Intelligent Systems(MFI), International Conference on, pages 1–6. IEEE, 2014
15. https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence_natural_language_processing.htm