Filtering Sentiment from Social Media

Priyanka Aher[1], Manisha Yerandekar[2], Pooja Tidme[3], Deepali Ganjave[4],

[1] *Department of Computer Engineering, PVGCOE, Nashik*
[2] *Department of Computer Engineering, PVGCOE, Nashik*
[3]*Department of Computer Engineering, PVGCOE, Nashik*
[4] *Department of Computer Engineering, PVGCOE, Nashik*

ABSTRACT

Now a days social media is playing a important role in affecting people's feelings in side or opposite a government or an institution. Therefore, to understand the sentiment of any posting in social media, an effective procedure is necessity. We have analyzed some Facebook postings to understand political sentiments. In any politically infected posting, there are some dominant words. At beginning, we have created a dictionary consisting of unique words gathered from political or nonpolitical posts or comments and then trained using Naïve Bayes algorithm based on probability algorithm. To identify the sentiment expressed in a new post or comment, we have taken each word of the posting and then compared those with the dictionary words for classification. At last, we have tested our algorithm using two hundred postings from Facebook and our result shows that the method can classify posts or comments with good accuracy. The field "political marketing", although relatively new, has grown quickly over the last few years. Now, it attracts scholars from a number of disciplines outside the mainstream marketing field. Political parties started to use the marketing instruments as their electoral campaigning. The term "Political marketing" contain these activities made by political parties to affect voters and is aimed on affecting the individuals with respect to political candidates to reach the maximum number of vote. One of the marketing instruments that political parties use to reach the voters is the use of social media. Since the increase of the internet in the early 1990s, the world's networked population has increased from the millions to the billions. Social media have become very important for life for society all over.. As the communications landscape gets more complicated, and more participatory, the networked population is getting much more importance and gaining greater power since social media gives people, the opportunity to access information much more simply, and it also gives more opportunities to engage in public conversation. As a result, it increased the capability of people to undertake collective action. This paper checks the effects of political marketing and specially using social media on voters of the 30 March 2014 local elections in Turkey. At beginning, it reviews the theory of marketing in brief terms and political marketing as part of it. And then, the new and popular social media channels were identified

Keyword : - *Track post, sentiment analysis, Naïve Bayes algorithm, political words, nonpolitical words, classification, Facebook*

1. INTRODUCTION :

Due to the fast growth of the Internet, the networked population has been increasing to a few billions .Hence, social media plays a important role in the life of our society all over and people are getting better access to data, more opportunities to busy themselves in dealing public feedback and an enhanced capability to undertake collective action. In previous days of social networking, sites were viewed only for friendship or dating purposes. However, with the progress of time the characteristics of most of the social media are recently changing. For example, in 2008 and 2012 US presidential elections, social media were used for election campaign and for the enrollment of the youth. Again, in 2009, the politically interested people and the politicians between German general election to conduct the election used social media -twitter efficiently and to discuss political events besides, everyone knows about the improving power of social media during the Arab uprising in 2011. A lot of political peoples, academics and journalists have improved social media as an great force for good, claiming that democracy is just a tweet away. Moreover, political experienced believe that if people want to liberate a society, they need to use Internet in a proper

way. Internet has been known as an effective way of the weak person against their authoritarian motivators. It is now an popular phenomenon that on one side government thugs firing bullet and on the other side Considering the fast growth of micro-blogging, many researchers have tried to analyze the sentiments of people either in favor or not and tried to understanding their opinions .Expressing political content in social media is improving largely .

2. RELATED WORK

We have given some sample page lists from where we have collected necessary posts or comments. We have considered these 10 well-known political pages. We have used 150 posts and 50 comments from political and nonpolitical topics from the political face book pages. These are used for our primary training set. We have identified their class level based on their sentiments political or nonpolitical. In an attempt to classify a new text sentiment, we need a filter that can be used to predict the probability of a class level based on learning. We have used Naïve Bayes classification method that is based on Bayes rule. For testing, we have extracted each word from the test text and matched with the dictionary words. Total political probability is found by multiplying individually matched words conditional political probability. Similarly, total nonpolitical probability is found by multiplying individually matched words conditional nonpolitical probability. Comparing these two probabilities, we can easily predict the class level of the text.

3. PROBLEM STATEMENT

To detect and prevent the malicious political sentiments or comments from Facebook post which plays a important role in influencing the people's and societies sentiments, we proposed the filtering the political sentiments of by using textual information. We have provided dictionary for this. The dictionary contains the words that may hurt someone's sentiment.

4. SYSTEM ARCHITECTURE

In proposed system, we taking post or comment from Face book then extracting each word using string analysis. Then we matching each extracted word with Dictionary words. We have considered 250 posts and comments randomly collected from 10 different pages and tried to predict the class level of some sample sentences. We have implemented conditional probability based on training. Every post or comment has two types of value, one for political and the other for nonpolitical probability. If the political probability is greater than the nonpolitical probability, then it is leveled as political post or comment, otherwise it will be considered as nonpolitical post or comment.
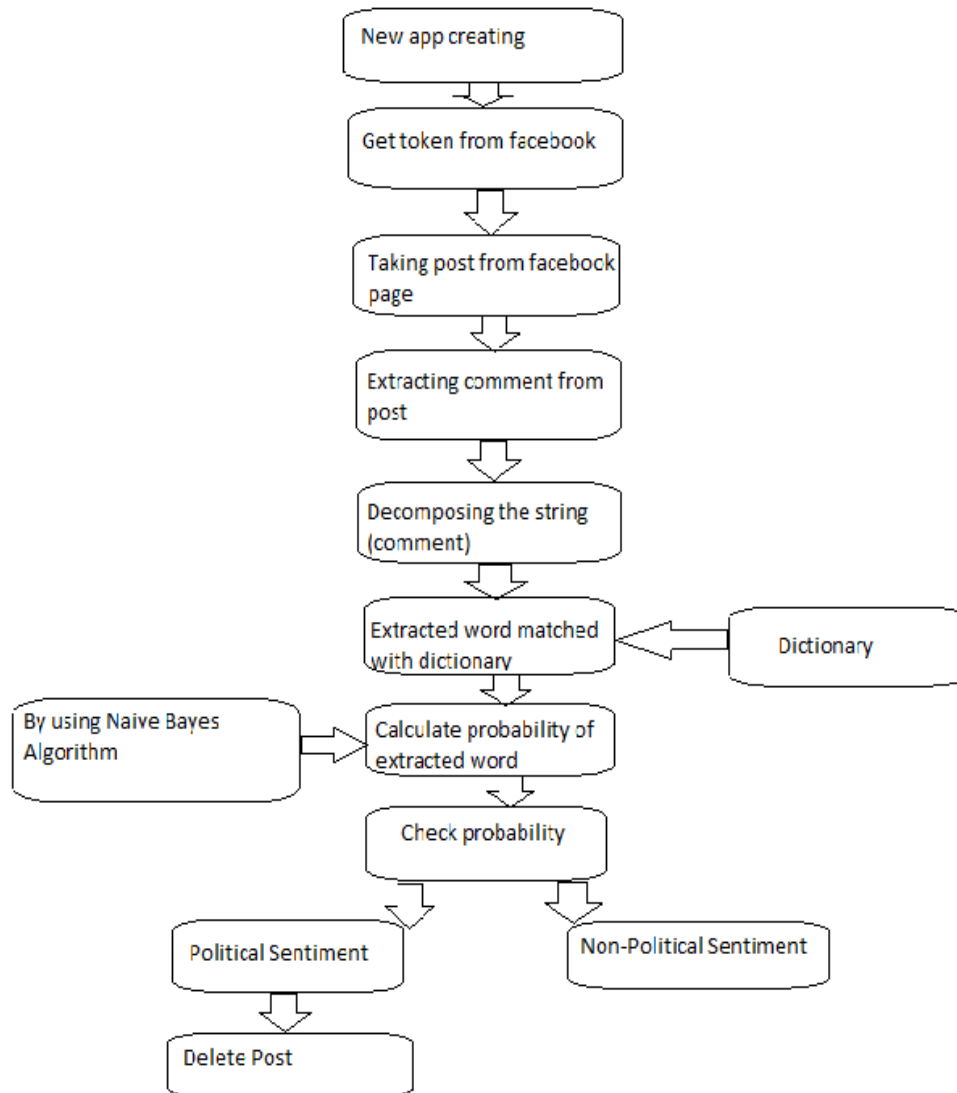
Fig -1: System Architecture

5. GOALS and OBJECTIVES

 Goals:
1. Time saving
2. Reliable
3. Accurate

Objectives:
1. To classify posts or comments with good accuracy.
2. To maintain peaceful environment in social media.
3. To reduce political turbulence.
4. To stop prank in social Media.

6. ALGORITHAM

Naive Bayes classifiers are a family of simple probabilistic classifiers based on applying Bayes' theorem with strong (naive) independence assumptions between the features. Naive Bayes has been studied extensively since the

1950s. It was introduced under a different name into the text retrieval community in the early 1960s, and remains a popular (baseline) method for text categorization, the problem of judging documents as belonging to one category or the other (such as spam or legitimate, sports or politics, etc.) with word frequencies as the features. With appropriate pre-processing, it is competitive in this domain with more advanced methods including support vector machines. It also finds application in automatic medical diagnose is.Naive Bayes classifiers are highly scalable, requiring a number of parameters linear in the number of variables (features/predictors) in a learning problem. Maximum-likelihood training can be done by evaluating a closed-form expression, which takes linear time, rather than by expensive iterative approximation as used for many other types of classifiers.

Algorithm Steps:

Step 1: Convert the data set into a frequency table

Step 2: Create state table by finding the probabilities

Step 3: Now, use Naive Bayesian equation to calculate the posterior probability for each class. The class with the highest posterior probability is the outcome of prediction.

Naive Bayesian Formula

$$P(c \mid x) = \frac{P(x \mid c) P(c)}{P(x)}$$

Likelihood — Class Prior Probability — Posterior Probability — Predictor Prior Probability

$$P(c \mid X) = P(x_1 \mid c) \times P(x_2 \mid c) \times \cdots \times P(x_n \mid c) \times P(c)$$

Fig 2: Naïve Bayesian Formula

7. SYSTEM OVERVIEW:

Facebook Comment And Post Tracker

User Login

Email ID

manisha1@gmail.com

Password

• • • • • • •

Login   Register

Unable To Login? Forgot Password ....

Fig 3: Login

Fig 4: Menu

Fig 5: Track Post

Fig 6: Post Detail

Fig 7: Comments Details

Fig 8: Detection
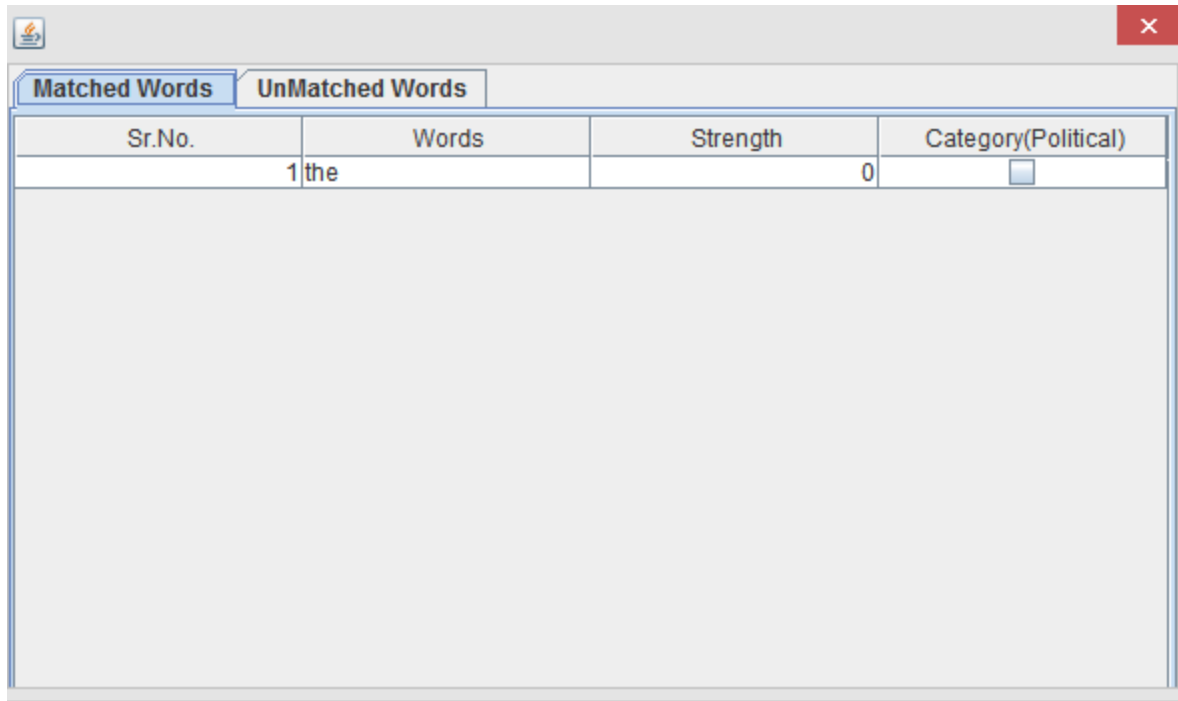
Fig 9: Matched words

8. CONCLUSIONS

Here filtering sentiment, using naïve Bayes is discussed. And the system is capable of filtering the political hazardous sentiments using text mining and web mining techniques. To understand sentiment of any social media post this method can be used.This can be used to create peaceful society by reducing political turbulence.The classification depends on dictionary that we provided and training to the words.

9. ACKNOWLEDGEMENT

10. REFERENCES

1. Elif Yolbulan Okan and Serhat Akyüz. The Role of Social Media in Political Marketing: 2014 Local Elections of Turkey European Journal of Business and Management, vol. 6, 2014.

2. Kristin Nicole Smith, Social media and political thesis, University of Tennessee, Knoxville, 2011.

3. Pascal Jürgens, Andreas Jungherr and Harald a Difference: New Gatekeepers and the _Filtering of Political Information on T ACM Web Sci 2011, Koblenz, Germany 2011.

4.The Arab uprising 2011: new media in the hands of a new generation in North Africa Aslib Proceedings, vol. 64, no.5, pp. 529-539, 2012. Journalism, vol. 12, no. 5, pp. 647-659, 2011.

5.Twitter as a Corpus for Sentiment Analysis and Opinion Mining, Language Resources and Evaluation Conference (LREc), vol. 10, pp.1320-1326, 2010.

6.Efthymios Twitter Sentiment Analysis: The Good the Bad and the OMG!, Icesm, vol. 11, pp. 538-541, 2011.

7. Apoorv Agarwal, Boyi Xie, Ilia Vovsha and Rebecca Passonneau, Sentiment Analysis of Twitter Data Proceedings of the workshop on languages in social media, Association for Computational Linguistics, pp. 33-38, 2011