

Sentiment Analysis on Political Tweets

Ayeena Malik¹, Divya Kapoor², Amit Prakash Singh³

¹Sri Guru Tegh Bahadur Institute of Management & Information Technology, Nanak

¹Gurdwara Nanak ²Piao Campus G.T. Karnal Road, Delhi-110033

³University School of ICT, GGS Indraprastha University, Dwarka, New Delhi

¹ayeenamalik072@gmail.com

Abstract: Sentiment analysis can be defined as a process of analyzing user reviews, opinions, emotions, sentiments, attitude regarding various entities such as products, services, organization, key issues. The Web is place where anybody can express their opinions or post reviews about various entities. Social Media is actually influencing the user's point of view or individual opinions. It has now become an essential part of digital marketing. Sentiment Analysis is a process of identifying and aggregating opinions of various about any issues or products. Monitoring Social Media activities has now become a prime concern for politicians to understanding their social image. In this paper we are going to analyze the tweets of various users regarding a political leader, which showcases the responses or opinions of voters pre and post-elections.

Keywords: Sentiment Analysis, political tweets, social media, opinions

1. INTRODUCTION

Sentiment analysis can be defined as a process of analysing user reviews, opinions, emotions, sentiments, attitude regarding various entities such as products, services, organization, key issues. The Web is place where anybody can express their opinions or post reviews about various entities. Social Media is actually influencing the user's point of view or individual opinions. It has now become an essential part of digital marketing. Sentiment Analysis is a process of identifying and aggregating opinions of various about any issues or products. Sentiment analysis is similar to opinion mining, with the help of which we can analyze the user's view by detecting the polarity of the text. Polarity means we can analyze the text on the basis of their views posted online i.e. to determine whether the piece of writing is positive, negative or neutral [1-5]. Sentiment Analysis is used to discover how people feel about a particular topic or a product. To buy a product the user checks the various social networking sites for reviews. These reviews and ratings on the Internet matters a lot to evaluate the product. Monitoring social media activities is necessary to evaluate the customer's loyalty and also to keep track on their sentiments regarding a brand or a product, impact of various marketing activities and many other aspects.

Using Natural language processing, statistics or machine learning methods to extract identify and characterize the sentiment content of a text unit[6]. Textual information in the world can be broadly categorized into two main types: facts and opinions. Facts are objective expressions about entities, events and their properties. Opinions are usually subjective expressions that describe people's sentiments, appraisals or feelings toward entities, events and their properties.

2. LITERATURE REVIEW

Fedrico Neri [8] describes a sentiment analysis study performed on about 1000 Face book posts about new casts and comparing the sentiments of Rai.In which two techniques of Sentiment Analysis can be used:

Supervised Machine Learning- It requires a training set of texts with manually assigned polarity values for eg. they learn features(words) that correlates with the value.

Unsupervised Machine Learning- That uses a lexicon with words scored for polarity values such as neutral, positive or negative. To extract opinions and their sentiments the following logical components are there:

- Crawler-Gathers document from internet
- Semantic engine –identifies relevant knowledge in the texts
- Search engine- enables natural language, semantic and semantic role queries.
- Machine translation engine- enables automatic translation of search results.
- Geo-referential engine- enables an interactive geographical representation of documents.
- Classification engine-classifies search results into clusters and sub-clusters, highlighting meaningful information.

Yasufumi Takama [9] in their paper proposed method by which a user profile is generated by using user's TV watching behavior using sentiment analysis. The method generates positive and negative profile with bookmark

format for each user based on iEPG and estimated rating of watching TV programme. If user watches program for long time that means he/she is interested and if changes channel frequently that means he/she is not interested. Daniel E.O' Leary in his paper focused on blog mining-reviews. He termed blog as type of media that allows user to present a range of information including opinions and personal views. Organizations are looking upon blogs as an important source of information and knowledge. Blogs may contain redundant and co related information about any topic financial, political, entertainment and news. It represents a single individual or a group. Blog mining is a process of searching and analyzing blogs in order to generate additional insights that might otherwise not be found by examining a single blog [8]. Capturing information from blogs is very useful to various marketing researches about the sentiments of buyers and potential buyers of their products. Sometimes sample of blogs are selected when the organization is concerned with opinions of a particular group or community. Random sample of blogs are selected when we require a broad base of opinions. Particular type of all available blogs are selected when we require an opinion according to a particular context. Blogs can be selected on the bases of time or a particular topic. Resources of a blog can be Internal if generated by employees of the organization or can be External if taken from outside sources i.e. customers. Opinions and sentiments can be determined with the help of frequency of appearance of words. To analyze the frequency of appearance of words we look for particular words or phrases used. Almost 50% of the blogs are Splogs i.e. Spam Blogs and are designed to induce readers to click advertisement and to get search engine to capture number of links. Polarity of the blogs can be examined by the tags placed on the blogs. The unit of reference may be word level, sentence level, or paragraph level and with the help of mood declaration we may classify the blogs as positive, negative or neutral. Blogs provide a platform for bloggers to express their opinions on a range of issues.

Namrata Godbole [9] proposed the large scale sentiment analyses for news and blogs. The frequency of adjectives is tracked with positive and negative polarity using WordNet. Sentiment hop counts to determine the strength of the candidate terms and eliminate ambiguous terms. Adjectives separated by "And" have same polarity but those separated by "But" have opposite polarity. Machine learning techniques perform better than simple counting methods and are more accurate.

3. METHODOLOGY USED

This paper examines the sentiment analysis of a prominent political figure by analysing a dataset gathered from the micro blogging site that is twitter. The dataset is obtained from twitter during the Prime Minister Narendra Modi's visit to United States of America. The project work starts with the creation of a twitter application which helps us

fetch certain access keys and tokens for the extraction and filtering of the tweets as required by the user and putting it in a database for further mapping of these tweets for sentiment analysis of the same. The access keys are generated and the tweets are extracted and put in the database. The approach to extract the tweets from the twitter database according to the user's need is a generic one as the user has the control over how many and which tweets are required by the user and he/she can work according to the desired level of complexity. The initial steps to extract tweets from a twitter account which is accomplished by making an account on twitter. From the respective account, a twitter development application is created on the developers end through an application on twitter site which gives the user the access to certain keys. These keys are used in the code to extract the tweets from the database and save them in a .csv file. The .csv file generated is used as an input for the next modules in the code to generate the required result from the analysis of the sentiments. With the grant of the keys, now the extraction of the tweets starts. The access tokens are now generated via the keys. With the help of tokenization, the tweets are split to words and every word is referred with a dictionary and its polarity is determined. We are using dictionary i.e. AFFINN-111. We assign a sentiment score to the words (S_w) to those words that are not present in AFINN-111 through this simple formula:

$$S_w = \frac{\sum_1^n S_{t(w \in t)}}{n}$$

where,

N is the number of tweets that contain the word

S_t is the sentiment score of the tweet that contains the word W .

The sentiment of a tweet is determined based on the higher value of s_+ or s_- .

If the values are equal the tweets are considered neutral.

$$s_t = \begin{cases} \text{positive} & \text{if } s_+ > s_- \\ \text{negative} & \text{if } s_+ \leq s_- \end{cases}$$

4. RESULTS

In this paper we are analyzing the tweets of our PM Narendra Modi before and after his foreign visits. The results are presented in the table 1:

TABLE 1 : Polarity score of tweets PM Narendra Modi

Polarity score	Modi	Modi After Foreign visits
Positive	21	27
Negative	-19	-2
Neutral	14	12

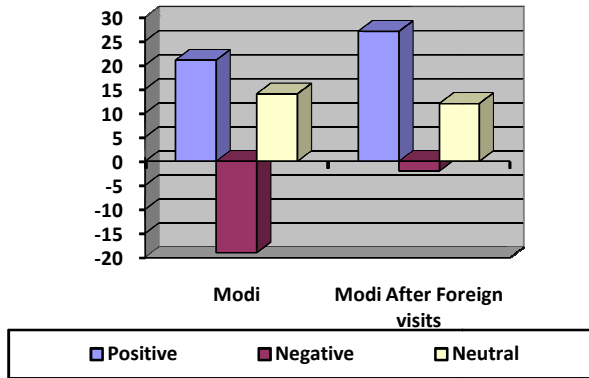


Fig. 1. Polarity score of tweets PM Narendra Modi

5. CONCLUSIONS

The use of sentiment analysis for tweets has a wide scope as it analyzes the sentiment of the people in an area and thus all the strategies and estimation can be based on how the people of that area respond to various political mandates. The sentiment analysis can be very useful to identify the grey areas for a political leader and thus help the respective party to ennoble the status of the leader so as to win the elections. It can also be a useful tool in predicting the results of an election with high accuracy and thus provide help to take the required action accordingly. All in all, with such a tool to judge the preferences of the people in terms of politics could hold the key to a leader's win in an election and also could help improve the strategies that are to be taken up by the party so that they are in best interest of the people being based on people's own preference. With such a tool in hand, the tide could be completing reversed and that matters a lot in politics. Thus, sentiment analysis may hold the key to the political decisions in the times to come.

6. FUTURE SCOPE

The data available online is inevitable which can be further useful in analyzing sentiments of people online. Sentiment analysis or opinion mining finds its application in almost every field. Most brands are using sentiment analysis tools to view and analyze the opinions of the actual users who are using their brand products. Similar to this the brands can also have an idea of what kind of services are appreciated by the customers along with their products. With the help of opinion mining news and media houses are trying to analyze the image of political leaders either after or before elections. But still a lot of challenges come in the way of sentiment analysis as there are some cases when the reviews available are influenced or paid reviews. The future of sentiment analysis lies in not only attaining accurate results but also detecting fake reviews

REFERENCES

[1] Liu, B.; Indurkha, N.; Damerau. Sentiment Analysis and Subjectivity. Handbook of Natural Language Processing. 2010.

[2] Dave, K.; Lawrence, S.; Pennock. D.M. Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews. In Proceedings of the 12th International Conference on World Wide Web, 2013.

[3] Hu M. Mining and Summarizing Customer Reviews. In Proceedings of ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2004.

[4] Pang, B.; Lee, L. A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts. In Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics, 2004.

[5] Pang, B.; Lee, L.; Vaithyanathan, S. Thumbs Up? Sentiment Classification Using Machine Learning Techniques. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2002.

[6] Pang, B.; Lee, L. Opinion Mining and Sentiment Analysis. In Foundations and Trends in Information Retrieval, 2008.

[7] Mouthami, K.; Devi, K. N.; Bhaskaran V. M. Sentiment Analysis and Classification Based On Textual Reviews. In Proceeding 2013 of IEEE International Conference on Information Communication & Embedded System, 2013.

[8] Neri, F.; Aliprandi, C.; Capeci, F.; Caudros, M.; Tomas. Sentiment Analysis on Social Media. ACM International Conference on advances in Social Network Analysis and Mining, 2012.

[9] Takama, Y.; Moto, Y.; Profile Generation from TV Watching Behavior using Sentiment Analysis. ACM International Conferences on Web Intelligence and Agent Technology, 2007.

[10] Daniel, E. Blog mining-review and extensions. Decision Support Systems, Elsevier, 2011.

[11] Godbole, N.; Srinivasaiah, M.; Skiena, S. Large Scale Sentiment Analysis for News and Blogs. ICWSM, 2007.

[12] Wangg, H.; Liu, H.; Song, W.; Feature-Based Sentiment Analysis Approach for product reviews. Journal of Software, 2014.

[13] Medhat, W.; Hassan, A.; Korashy, H. Sentiment analysis algorithms and applications: A survey. Ain Shams Engineering Journal, 2014, 5, 4.

[14] Agarwal, A.; Xie, B.; Vovsha, I.; Rambow, O.; Passonneau, R. Sentiment Analysis of Twitter Data. ACM Proceeding of LSM 2011 of the workshop on Languages in Social Media, 2011, 30-38.

[15] Asur, S.; Huberman, B. A. Predicting the Future With Social Media. IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, 2010.