Sentiment Analysis and its Impact in Modeling Election Scenario

Dr. Gagandeep Jagdev

Dept. of Computer Science, Punjabi University Guru Kashi College, Damdama Sahib (PB).

*Corresponding Author: *Dr. Gagandeep Jagdev, Dept. of Computer Science, Punjabi University Guru Kashi College, Damdama Sahib (PB).*

Abstract: *Social media has emerged as a platform where general public can freely express their opinions, views, sentiments, and thoughts about any political leader or political party. Often some news agencies manipulate the news and delivers it in such a manner that it favors one political party and goes against other. But the sentiments expressed by masses on social media is the true image of any political leader or political party. This research paper if focused on elaborating about sentiments and opinions along with their types. The paper also discusses about classifying the keywords in any post or tweet as positive, negative, or neutral sentiment.*

Keywords: *opinions, sentiments, sentiment analysis, social media.*

1. INTRODUCTION

There is very common practice in India of people claiming that the news channel are always bias towards some political parties and have soft corner and favor other political parties. Different political parties accuse the mainstream news agencies of favoring their opposition, and to large extent it is true. Today, with the growth of internet access, general public is free to express their opinions related to elections and have a better understanding of prevailing political scenario. More and more people are posting their views, opinions, and sentiments about current political affairs through social networking sites, blogs, and microblogs. It was found that during Indian general election in the year 2014, in the time period of four months, the conversations having relevance with Indian elections were more than twice the conversations that took place during whole of the year 2013 and the twitter users account of Indians were also doubled [1, 2]. The detailed description of sentiments and opinions are discussed as under.

1.1. Sentiments

Sentiment analysis involves categorizing opinions in text into categories like "positive" or "negative" often with an implied category of "neutral" as shown in Fig. 1. Sentiment analysis is also called opinion mining or voice of the customer. Sentiment analysis is the computational job of automatically determining what feelings a writer is expressing in text [3, 4]. Sentiment is often framed as a binary distinction (positive vs. negative), but it can also be a more fine-grained, like identifying the specific emotion an author is expressing (like fear, joy or anger).Sentiment analysis is used for many applications, especially in business intelligence [5]. Some examples of applications for sentiment analysis include:

- Analyzing the social media discussion around a certain topic
- Evaluating survey responses
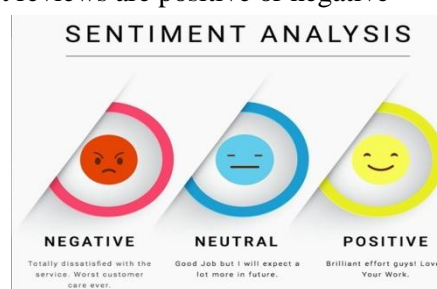- Determining whether product reviews are positive or negative

Fig1. *The figure depicts the classification of sentiments in three broad categories*

Some key points to be considered in sentiment analysis are mentioned as under [17, 18].

- Create or find a list of words associated with strongly positive or negative sentiment.

- Count the number of positive and negative words in the text.

- Analyze the mix of positive to negative words. Many positive words and few negative words indicates positive sentiment, while many negative words and few positive words indicates negative sentiment.

The initial step of lexicon, i.e. creating a word list is a bit tedious task and is time consuming. Often there is a need to modify the lexicon if one is discussing on a specific topic. The meaning of the one same word changes as per the sentence or situation in which it is used. For example, the word "sick" can have both positive and negative sentiment. If one is talking about a pet store which sells a lot of sick animals, then surely here the word "sick" reflect a negative sentiment. But if the same "sick" word is used for talking about a skateboarding instructor who teaches a lot of sick flips, the word reflects a positive sentiment [6, 7, 17].

Usually, besides identifying the opinion, these systems extract attributes of the expression e.g.:

- Polarity: if the speaker express a positive or negative opinion.
- Subject: the thing that is being talked about.
- Opinion holder: the person, or entity that expresses the opinion.

Sentiment analysis is a much popular topic and of great interest in several applications. As information on internet is constantly growing, many text expressing opinions are available on social media. This sentiment analysis is prominently used in product reviews, marketing analysis, public relations, net promoter scoring, product feedback, and customer service [8, 9, 18].
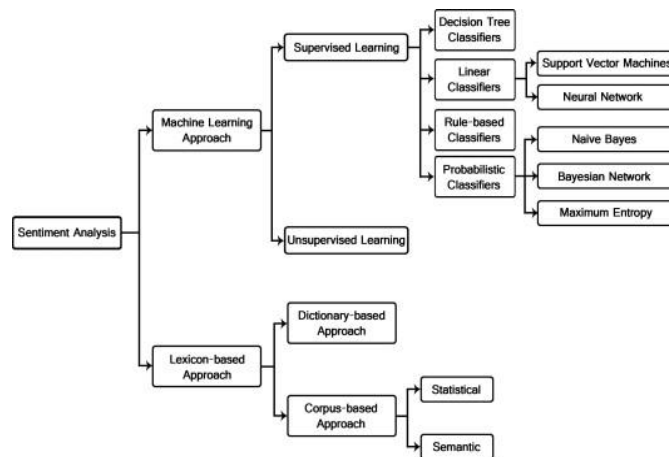

Fig2. *The figure depicts the different approaches involved in sentiment analysis.*

1.2. Opinions

Broadly text information is divided into two categories, facts and opinions. Facts are objective expressions and opinions are usually subjective expressions relevant to people appraisals, sentiments, and feelings towards ant subject of interest. When one classifies a sentence as subjective or objective, it is referred as subjectivity classification. When sentence is classified as positive, negative or neutral opinion, it is referred as polarity classification [10, 11, 17].

Types of opinions

The opinions are broadly classified into below mentioned categories.

1.2.1. Direct Opinion Vs. Comparative Opinions

- Direct opinion provides an opinion directly about an entity. For instance: "The picture quality of camera X is poor". This sentence directly gives a negative opinion about camera X.
- In case of comparative opinion, the opinion is conveyed by performing comparison of entity with another. For instance: "The picture quality of camera X is better than that of camera Y". The comparative opinion conveys differences or similarities between two or more entities by making

use of superlative form of an adverb or an adjective. The sentence under study establishes a positive opinion about camera X and negative opinion about camera Y.

1.2.2. Explicit Vs. Implicit Opinions

- An opinion explicitly expressed in a subjective sentence is referred as explicit opinion. For instance, the below mentioned sentence expresses an explicit positive opinion. "The voice quality of this phone is amazing".

- An opinion implied or indirect in an objective sentence is called an implicit opinion. For instance, the below mentioned sentence expresses an implicit negative opinion. "The earphone broke in two days."

1.3. Different Levels of Analysis

The research relevant to sentiment analysis is carried out at three different levels of granularity. The levels are briefly mentioned as under [12, 13].

Document level – At document level, it is classified whether the whole document articulates a positive sentiment or a negative sentiment.

Sentiment analysis research has been mainly carried out at three levels of granularity: document level, sentence level, and aspect level. We briefly introduce them here.

Document level. The task at the document level is to classify whether a whole opinion document expresses a positive or negative sentiment (Pang et al., 2002; Turney, 2002). It is thus known as document-level sentiment classification. For example, given a product review, the system determines whether the review expresses an overall positive or negative opinion about the product. This level of analysis implicitly assumes that each document expresses opinions on a single entity (e.g., a single product or service). Thus it is not applicable to documents that evaluate or compare multiple entities, for which more fine-grained analysis is needed. We study document-level sentiment analysis in Chapter 3.

2. LITERATURE REVIEW

Opinions are very significant to businesses and organizations because they always want to find consumer or public opinions about their products and services. Local and federal governments also want to know public opinions about their existing or proposed policies. Such opinions will enable relevant government decision makers to respond quickly to the fast-changing social, economic, and political climates. In international politics, every government wants to monitor the social media of other countries to find out what is happening in these countries and what people's views and sentiments are about current local and international issues and events. Such information is very useful to diplomacy, international relations, and economic decision making. Similarly, the opinions matter when it comes to elections. Each political leader and every political party wants to know that what people thinks about their political performance, what kind of status they enjoy among general public, where does common ma evaluates them, and what else is expected and what was not expected from them.

Several researchers have also investigated sentiments of public opinions in the context of electoral politics. For example, in O'Connor et al. (2010), a sentiment score was computed based simply on counting positive and negative sentiment words, which was shown to correlate well with presidential approval, political election polls, and consumer confidence surveys. In Bermingham and Smeaton (2011), tweet volume and positive and negative tweets were utilized as the independent variables and polling results as values for the dependent variable to train a linear regression model to predict election results. In Chung and Mustafaraj (2011) and Gayo-Avello et al. (2011), several limitations of current works on using Twitter data to predict political elections were discussed, one of them being poor sentiment analysis accuracy. The works in Diakopoulos and Shamma (2010) and Sang and Bos (2012) used manually annotated sentiments of tweets for election prediction. Tumasjan et al. (2010) even showed that simple party mentions on Twitter can be a good predictor of election results. In other related works, Yano and Smith (2010) reported a method for predicting comment volumes of political blogs, Chen et al. (2010) studied political standpoints, and Khoo et al. (2012) analyzed sentiment in political news articles about economic policies and political figures.

3. INFLUENCE OF SOCIAL MEDIA ON ELECTIONS

The impact of social media on general public in regard with elections is very deep. Fig. 3 shows the logos of the famous social media platforms [14, 15, 17].

Fig3. *The figure shows the logos of popular social media sources*

The influence made by social media are briefed as follows:

3.1. Voter Behavior Will be Better Predicted

The social networks has proved highly valuable for general voters when it comes to gain information about any political leader or political party. The user's post act as source to gain knowledge about voter's preferences, the issues influencing decision making. The voting pool can be better understood by the user's news feed.

3.2. Candidates Using Social Media More Personally

The political leaders have started using their accounts for creating one-on-one connection. The political territory has caught up with this. The political leaders can use this as a tool for creating powerful connections between them and their voters to initiate interest and positively affect voter turnout.

3.3. Diversified Social Media Channels

An evaluation of Facebook data during the campaign showed that more than 95% of buzz around a candidate were determined by posts outside of that candidate's primary campaign page. This cycle saw the rise of a more distributed social media presence for both sides. Each candidate had their own primary Facebook page, but many other pages were also formed and sustained for family members, campaign substitutes, and the National Committees of both parties. In the future, we should see more of this expanded but interrelated approach to social media campaigning. Fig. 4 shows the glimpse of the news channel depicting sentiment comparison of different political leader among masses on twitter [16, 18].

Fig4. *The figure shows the glimpse of the news channel depicting sentiment comparison of different political leader among masses on twitter*

4. CLASSIFICATION OF SENTIMENT RELEVANT WORDS

Fig. 5 shows the flowchart describing the process followed for creating the dictionaries having positive, negative, and neutral sentiments.
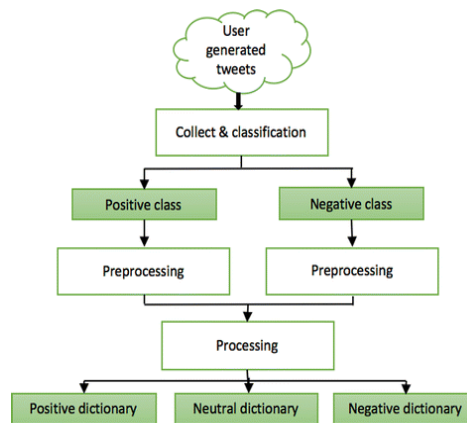
Fig5. *The figure shows the flowchart describing the process followed for creating the dictionaries having positive, negative, and neutral sentiments*

Algorithm for dictionaries creation

- Firstly, the user tweets are generated.
- These tweets are than collected and classified broadly into positive class and negative class.
- This step is followed by performing the preprocessing of both the classes and obtaining processed and relevant data.
- Finally, the processed data is further classified into three separate dictionaries: positive dictionary, negative dictionary, and neutral dictionary.

5. CONCLUSION

Social media has evolved as a major source for general public to express their opinions and sentiments. Election results can be fairly predicted by analyzing and studying the people posts, sentiments, and chats on social networking sites like Facebook and Twitter. The research paper discussed the about sentiments and its type. The paper also elaborated about opinions and its types. The classification of sentiments done in the paper can be effectively used to create appropriate dictionaries and further analyzing the posts and tweets of general public to know about their views regarding any upcoming elections.

REFERENCES

[1] B. Agarwal, N. Mittal, "Prominent Feature Extraction for Review Analysis: An Empirical Study", In Journal of Experimental and theoretical Artificial Intelli-gence, 2014, DOI: 10.1080/0952813X.2014.977830.

[2] S Bandyopadhyay and K Mallick, "A New Path Based Hybrid Measure for Gene Ontology Similarity", IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol.11, no. 1, pp. 116-127, Jan.-Feb. 2014, doi:10.1109/TCBB.2013.149

[3] Mariana Romanyshyn(2013). Rule-Based Sentiment Analysis of Ukrainian Re-views. International Journal of Artificial Intelligence & Applications (IJAIA), Vol. 4, No. 4, July 2013

[4] Gagandeep Jagdev et al., "A Comparative study of Conventional Data Mining Algorithms against Map-Reduce Algorithm", in International Journal of Advance Research in Science and Engineering (IJARSE), ISSN (O) – 2319-8354, ISSN (P) – 2319-8346,Volume – 06, Issue – 05, May 2017.

[5] Gagandeep Jagdev et al., "Analyzing Maneuver of Hadoop Framework and MapR Algorithm Proficient in supervising Big Data", in International Journal of Advanced Technology in Engineering and Science (IJATES), ISSN – 2348-7550,Volume – 05, Issue – 05, May 2017.

[6] Jose, R. and Chooralil, V. S. (2015). Prediction of election result by enhanced sentimentanalysis on twitter data using word sense disambiguation. In Control Communication& Computing India (ICCC), 2015 International Conference on, pages 638–641. IEEE.

[7] Gagandeep Jagdev et.al (2017). "Exploring Application of Big Data in Elections – From Data to Action", International Journal of Research Studies in Computer Science and Engineering (IJRSCSE), ISSN: 2349-4840 (P), ISSN: 2349-4859(O), Volume 4, Issue 4, pp.64-71, DOI: http://dx.doi.org/10.20431/2349-4859.0404008.

[8] Gagandeep Jagdev et al., "Analyzing Maneuver of Hadoop Framework and MapR Algorithm Proficient in supervising Big Data", in International Journal of Advanced Technology in Engineering and Science (IJATES), ISSN – 2348-7550,Volume – 05, Issue – 05, May 2017.

[9] Padmaja, S., Fatima, S. S., and Bandu, S. (2013). Analysis of sentiment on newspaperquotations: A preliminary experiment. In Computing, Communications and NetworkingTechnologies (ICCCNT), 2013 Fourth International Conference on, pages 1–5.IEEE.

[10] Gagandeep Jagdev et al., "A Study of Clustering and Classification Techniques involved in Data Mining", in International Journal of Advanced Technology in Engineering and Science (IJATES), ISSN – 2348-7550,Volume – 05, Issue – 05, May 2017.

[11] Gagandeep Jagdev et al., "Analyzing and Filtering Big Data concerned with elections via Hadoop Framework" in International Journal of Advance Research in Science and Engineering (IJARSE), ISSN (O): 2319-8354, ISSN (P): 2319-8346, Volume No. 6, Issue No, 4, April 2017.

[12] Sharma, A. and Dey, S. (2012). A comparative study of feature selection and machinelearning techniques for sentiment analysis. In Proceedings of the 2012 ACM Researchin Applied Computation Symposium, pages 1–7. ACM.

[13] Gagandeep Jagdev et al., "Comparing Conventional Data Mining Algorithms with Hadoop based Map-Reduce Algorithm considering elections perspective", in International Journal of Innovative Research in Science and Engineering (IJIRSE), ISSN: 2454-9665 (O), ISSN: 2455-0663(P), Volume – 3, Issue – 3, March 2017.

[14] Gagandeep Jagdev et al., "Big Data Proposes an Innovative concept for contesting elections in Indian Subcontinent" in International Journal of Scientific and Technical Advancements (IJSTA), ISSN-2454-1532, 2015.

[15] Gagandeep Jagdev et al., "Excavating Big Data associated to Indian Elections Scenario via Apache Hadoop" in International Journal of Advanced Research in Computer Science (IJARCS), ISSN 0976 – 5697.

[16] Tumitan, D. and Becker, K. (2014). Sentiment-based features for predicting electionpolls: a case study on the brazilian scenario. In Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2014 IEEE/WIC/ACM International Joint Conferences on, volume 2, pages 126–133. IEEE.

[17] https://monkeylearn.com/sentiment-analysis/; last accessed: 20 September 2018.

[18] http://blog.kaggle.com/2017/10/05/data-science-101-sentiment-analysis-in-r-tutorial/; last accessed: 20 September 2018.

AUTHORS' BIOGRAPHY

Dr. Gagandeep Jagdev, is a faculty member in Dept. of Computer Science, Punjabi University Guru Kashi College, Damdama Sahib (PB). His total teaching experience is more than 12 years and has 125 international and national publications in reputed journals and conferences to his credit. He is also a member of editorial board of several international peer-reviewed journals and has been active Technical Program Committee member of several international and national conferences conducted by renowned universities and academic institutions. His field of expertise is Big Data, Data Mining, Image processing in Diabetic Retinopathy, ANN, Biometrics, RFID, Cloud Computing, Cloud security, Cryptography, and VANETS.