

Opinion Mining and Emotional Intelligence: Techniques and Methodology

B.Asraf yasmin¹, Dr.R.Latha²

¹ Ph.D Research Scholar, Computer Applications, St.Peter's University, Chennai.

² Prof & Head., Dept of Computer Applications, St.Peter's University, Chennai.

ABSTRACT

Humans communication is generally under the control of emotions and full of opinions. Emotions and their opinions plays an important role in thinking process of mind, influences the human actions too. Sentiment analysis is one of the ways to explore user's opinion made on any social media and networking site for various commercial applications in number of fields. This paper takes into account the basis requirements of opinion mining to explore the present techniques used to developed an full fledge system. Is highlights the opportunities or deployment and research of such systems. The available tools used for building such applications have even presented with their merits and limitations.

Keywords: Opinion Mining, Emotion, Sentiment Analysis, EM Algorithm, SVM algorithm.

INTRODUCTION

Emotions are the complex state of feelings that results in physical and emotional changes that influences our behavior. Emotion is a subjective, conscious experience characterized mainly by psycho-physiological expressions, biological reactions, and mental states. Emotion is often associated and considered commonly significant with mood, nature, personality, disposition, and motivation [1]. Emotion is a positive or negative experience that is associated with a particular pattern of physiological activity. Humans carry lot of emotions like happiness, sadness, angry, disgust, surprise, fear, panic, scared etc. identifying these emotions are very easy in face to face communication compare to written communication. But now a day's use of social media has increased rapidly and the huge amount of textual data became available on web, mining and managing this vast data has become a crucial task. As the growth of E-facilities have increased lots of people got encouraged to write their emotions, views, opinions about a person, product, place or anything they want. Opinion Mining or Sentiment analysis involves building a system to explore user's opinions made in blog posts, comments, reviews or tweets, about the product, policy or a topic [2].

Opinion mining is nothing but finding the opinion of person from sentences and classify them on the basis of polarity. As the world changed into E-World the way of expression has dramatically changed for example wide use of smiley's and symbols can be seen as expression while texting. Social communication can be observed on internet and new term has been coined for various ways of communication like texting, twitting, posting etc.

people like to communicate with others through internet, they want to share their feelings, likes, dislikes, opinions, views, reviews, emotions etc. people are happy to share their personal life via social media, the use of social media has increased so much and so rapidly that even no body worries about what they are sharing and is this good to share our personal life with unknown persons? Is there any need to share our photos, videos or our daily activities on internet? So finding the sentiment, emotion behind this activity is also an important task for understanding the psycho-socio status. So from that text, mining the opinions of people and finding their views, reaction, sentiments and emotions have become challenging task.

Opinion Mining is the field of study that analyzes people's opinion, sentiments, evaluations, attitudes and emotions from written text. Opinion Mining is one of the most active research areas in Natural Language Processing and is also widely studied in data mining, web mining and text mining this research has spread outside of computer science to the management science and social science due to its importance to business and society. The growing importance of sentiment analysis coincides with the growth of social media such as Reviews, Forums, discussion groups, chatting, blogs, micro blogs, twitter and social networks.

II. CATEGORIZATION OF TEXT

Sentiment analysis is also called as opinion mining; as it mines the information from various text forms such as reviews, news & blogs and classifies them on the basis of their polarity as positive, negative or neutral [3]. It focuses on categorizing the text at the level of subjective and objective nature. Subjectivity indicates that the text contains/bears opinion content for e.g. Battery life of Samsung mobiles is good. (This sentence has an opinion, it talks about the Samsung mobile phones and showing positive (good) opinion hence it is Subjective). Samsung mobiles are having long battery life. (This sentence is a fact, general information rather than an opinion or a view of some individual and hence its objective) [4]

III.COMPONENTS OF OPINION MINING

There are mainly three components of Opinion Mining [3]:

- **Opinion Holder:** Opinion holder is the holder of a particular opinion; it may be a person or an organization that holds the opinion. In the case of blogs and reviews, opinion holders are those persons who write these reviews or blogs.
- **Opinion Object:** Opinion object is an object on which the opinion holder is expressing the opinion.
- **Opinion Orientation:** Opinion orientation of an opinion on an object determines whether the opinion of an opinion holder about an object is positive, negative or neutral.

IV. DIFFERENT LEVELS OF SENTIMENT ANALYSIS

In general, sentiment analysis has been investigated mainly at three levels [4].

- **Document level:** The task at this level is to classify whether a whole opinion document expresses a positive or negative sentiment. For example, given a product review, the system determines whether the review expresses an overall positive or negative opinion about the product. This task is commonly known as document level sentiment classification.
- **Sentence level:** The task at this level goes to the sentences and determines whether each sentence expressed a positive, negative, or neutral opinion. Neutral usually means no opinion. This level of analysis is closely related to subjectivity classification which distinguishes sentences (called objective sentences) that express factual information from sentences (called subjective sentences) that express subjective views and opinions.
- **Entity and Aspect level:** Both the document-level and sentence-level analyses do not discover what exactly people liked and did not like. Aspect level performs fine-grained analysis. Aspect level was earlier called feature level (feature based opinion mining and summarization).

V. CHALLENGES & METHODOLOGY IN OPINION MINING

There are several challenges in Opinion Mining as follows,

- **Domain-independence:** The biggest challenge faced by opinion mining and sentiment analysis is the domain dependent nature of sentiment words. One features set may give very good performance in one domain, at the same time it perform very poor in some other domain [5].
- **Asymmetry in availability of opinion mining software:** The opinion mining software is very expensive and currently affordable only to big organizations and government. It is beyond the common citizen's expectation. This should be available to all people, so that everyone gets benefit from it[6].
- **Detection of spam and fake reviews:** The web contains both authentic and spam contents. For effective Sentiment classification, this spam content should be eliminated before processing. This can be done by identifying duplicates, by detecting outliers and by considering reputation of reviewer[5].
- **Incorporation of opinion with implicit and behavior data:** For successful analysis of sentiment, the opinion words should integrate with implicit data. The implicit data determine the actual behavior of sentiment words [6].
- **Mixed Sentences:** Suppose the word is positive in one situation may be negative in another situation. For e.g. Word LONG, suppose if customer says the battery life of Samsung mobile is too long so that would be a positive opinion. But suppose if customer says that Samsung mobile take too long time to start or to charge so it would be a negative opinion.
- **Way of Expressing the Opinion:** The people don't always express opinions in the same way. The opinion of every individual is different because the way of thinking, the way of expressing is vary from person to person.



- **Use of Abbreviations and short forms:** People using social media more and that to for chatting, expressing their views using shortcuts or abbreviations so the use of colloquial words is increased. Uses of abbreviation, synonyms, special symbols is also increase day by day so finding opinion from that is too difficult. For e.g. F9 for fine, thnx for thanks, u for you, b4 for before, b'coz for because, h r u for how are you etc.
- **Typographical Errors:** Sometimes typographical errors cause problems while extracting opinions.
- **Orthographic Words:** People use orthographic words for expressing their excitement, happiness for e.g. Word Sooo..... Sweeetttt....., I am toooo Haappy or if they in hurry they stress the words for e.g. comeeeee fasssst I am waittttnggg.
- **Natural language processing overheads:** The natural language overhead like ambiguity, co-reference, Implicitness, inference etc. created hindrance in sentiment analysis tool [6].

VI. DATA SOURCES AND TOOLS OF OPINION MINING

While doing research the collection of data is the biggest issue and for the task like opinion mining, sentiment analysis it's too difficult because lots of information is available on internet and collection of that data and extraction of opinion from huge amount of data is too hard. So here we discussed about some available data sources and the tools which is used for extraction the sentiments and opinion of the given text.

6.1. Data sources available for opinion mining

There are various data sources available on web, i.e. Blogs, Micro blogs, online posts, News feeds, Forums, review sites etc.

- **Blogs:** Blogs are nothing but the user own space or diary on internet where they can share their views, opinions about topics they want.
- **Online Reviews:** on Internet various review sites are available through that you can check online reviews of any product before purchasing that.
- **Micro blogging:** Micro blogs allow users to exchange small elements of content such as short sentences, individual images, or video links", which may be the major reason for their popularity.
- **Online Posts:** people share their own ideas, opinions, photos, videos, views, likes, dislikes, comments on specific topics etc.
- **Forums:** An Internet forum, or message board, is an online discussion site where people can hold conversations in the form of posted messages.

6.2. Classification methods in opinion mining

- **Corpus Based Approach:** A popular corpus-driven method is determining words emotional affinity to learn their probabilistic affective scores from a large corpus. The method is to assign a happiness factor to words based on their occurrence frequency in happy labeled blog posts compared to total frequency in a corpus of blog

posts labeled with happy and sad mood annotations. They compare happiness factor of scores of words with scores in the list.

- **Dictionary Based Approach:** Dictionary based approach uses lexical resources like Word Net to automatically obtain emotion-related words for emotion classification. Beginning from primary emotion adjectives set, they retrieve similar words from Word Net using senses of all words in synsets with emotion adjectives. The process exploits synonym and hyponym relations in Word Net to manually locate words similar to nominal emotion words. Affective weights are acquired automatically from a very large text corpus in an unsupervised way.
- **Feature based opinion mining:** Using OM, a review is evaluated at 3 levels- document, sentence and feature levels. Evaluating a review at document level, the entire review is classified positive or negative based on opinions expressed in that review. When evaluated at sentence level, then every sentence is classified as positive or negative while feature level or feature based OM gives summary where a product feature is liked or disliked by reviewers. The major feature based OM tasks include - (1) identifying products features in review, (2) determining opinion expressed by reviewer (positive, negative or neutral), (3) summarizing discovered information.

VII. CONCLUSION

Emotions are often associated and considered commonly significant with mood, nature, personality, disposition, and motivation. Opinion Mining or Sentiment analysis refers to extraction of opinion from given text and classify them on the basis of polarity i.e. positive, negative and neutral. In this paper, we discussed about various levels of sentiment analysis and technique used to identify and extract opinions. Here we gave some challenges used while working on opinion mining like some orthographic errors, typographical mistakes, abbreviations, colloquial words etc. are the major challenges. This paper provides a brief review to cover the major challenges, stages, application and advantages of opinion mining. In our study, we find some techniques like Naïve Bayes, Maximum Entropy, and SVM etc. are very often used in opinion mining and sentiment analysis.

REFERENCES

- [1] Cacioppo, 27 Nov 2014 “Studyguide for DiscoveringPsychology: The Science of Mind,” John Publication.
- [2] S. ChandraKala and C. Sindhu, October 2012 “Opinion Mining and Sentiment Classification: A Survey,” ICTACT Journal On Soft Computing, Volume: 03, Issue: 01.
- [3] Sharma .R. et al., March 2014 “Opinion Mining in Hindi Language: A Survey,” International Journal in Computer Science & Technology (IJFCST), Vol 4, No. 2.
- [4] Bing L., 2012 “Sentiment Analysis and Opinion Mining,” Morgan & Claypool Publishers.
- [5] Liu, B., 2010 “Sentiment Analysis and Subjectivity,” Appeared in Handbook of Natural Language Processing, Indurkha, N. & Damerau, F.J. [Eds.].

- [6] Haseena Rahmath, 2014 “Opinion Mining and Sentiment Analysis - Challenges and Applications,” International Journal of Application or Innovation in Engineering & Management (IJAIEM), Volume 3, Issue 5.
- [7] B. Pang, L. Lee, and S. Vaithyanathan, 2002 “Thumbs up? Sentiment classification using machine learning techniques,” Proceedings of the Conference on Empirical Methods in Natural Language Processing, pp. 79-86.
- [8] Turney P.D., 2002 “Thumbs up or down? Semantic orientation applied to unsupervised classification of reviews”, ACL, p. 417-424.
- [9] M.G. Madden, 2002 “A new Bayesian network structure for classification tasks,” Proceedings of 13th Irish Conference on Artificial Intelligence & Cognitive Science, vol. 2464.
- [10] Hai Leong Chieu . et al., “Named Entity Recognition: A Maximum Entropy Approach Using Global Information,” http://www.cnts.ua.ac.be/conll2003/pdf/16063_chi.pdf
- [11] Turney P.D., Littman M.L., 2003 “Measuring praise and criticism: Inference of semantic orientation from association,” ACM TOIS 21(4) , p. 315-346
- [12] S.-M. Kim and E. Hovy, 2004 “Determining the sentiment of opinions,” Proceedings COLING-04, the Conference on Computational Linguistics, Geneva, Switzerland.
- [14] Jack G. Conrad and Frank Schilder, June 4-8, 2007 “Opinion Mining in Legal Blogs,” ACM, ICAIL '07.
- [15] Xiaowen Ding, Bing Liu, 2008 “A Holistic LexiconBased Approach to Opinion Mining,” ACM WSDM'08, February 11-12.
- [16] Alec Go, Richa Bhayani, Lei Huang, 2009 “Twitter sentiment classification using distant supervision,” CS224N Project Report, Stanford.