A Review of Opinion Mining and Analysis

Prof. M. S. Deshmukh, Prof. Ms. P. V. Dudhe, Prof. Ms. N. V. Kadam, Dr. N. N. Khalsa

Abstract- **The impact of social media platforms is huge while considering the volume of internet traffic and the content it generates. The Sentiment Analysis is a technique which is widely used in the natural language processing. It is used for determining the sentiments of a text. Opinion mining which is also known as sentiment analysis is a sub field of text mining. In this the main task is to extract views or opinions from contents generated by web users while using social media. Opinion Mining is mainly to discover the opinions from unstructured text by combining techniques from Computer Science and natural language processing.**

Keywords-**Opinion Mining, Sentiment Analysis, natural language processing**

I. INTRODUCTION

Internet in today's life is like a lifeline of every individual. Social media is flooded with various forums, blogs, video sharing websites, coding platforms and social networking platforms. The web content is generated by social media platforms, which can also be mentioned as social media data, is diverse, including text, images, audio and videos. Each one has a social media account and are posting various activities from their life on social media. The comment or opinion put on social media also are in huge volume and has to be managed regularly. All the comments put on social media are put in different languages and it's a herculean task to analyze the data which is in terabyte. The opinion needs to be analyzed for a better solution for certain problem in specific area. That's where there is need for opinion analysis, sentiment analysis.

The human language is complex. That is the reason it's necessary to teach a machine, how to analyze the various grammatical mistakes, slang language, cultural differences, misspellings and other nuance things written on social media, is a very difficult process. For a machine it is very difficult to understand the effect of context on tone.

II. LITERATURE SURVEY

Opinion mining deals with on polarity detection and sentiment analysis is related with emotion recognition. For opinion mining and sentiment analysis, data mining and processing techniques are used. Processing technique processes the natural language used for commenting and giving opinion. These techniques help to uncover, retrieve, and refine Opinion from the internet's vast textual information.

Now a days, opinion mining and sentiment analysis research are evolved in both technique sophistication and analysis depth. Bo Pang [1] with her colleagues classified complete documents based on overall positive or negative polarity. They also gave rating scores of reviews. These documents were mainly supervised and labeled samples done manually. Examples samples like a movie or reviews of product explicitly indicating an overall polarity of an opinion.

Liu [7] defines an opinion of a 5-tuple which contains the target, its attribute at which the opinion is directed, the sentiment contained in the opinion whether it is negative or positive or neutral, the opinion holder's name and the date when the opinion was given.

There are other approaches, like A. Neviarouskaya presented in [6], where he attempts to mention the emotional states such as ''anger'', ''fear'', ''joy'', or ''interest'' instead of just positive or negative. In this case, model given by [7] could be enriched by adding another element to the opinion tuple model to represent this information.

Zhu [8], proposed aspect based opinion polling which are free from the textual customers reviews. The various aspects that are related withterms used for identification of aspects were learnt using a multi-aspect bootstrapping method. Jeonghee Yi et al., proposed a Sentiment Analyzer for extracting opinions about a subject from online data documents [9]. Sentiment analyzer uses natural language processing techniques. Alekh Agarwal et al., proposed a machine learning method incorporating linguistic knowledge gathered through synonymy graphs, for effective opinion classification. This approach shows the degree of influence among relationships of documents have on their sentiment analysis. Michael et al., presented, a prototype system for mining topics and sentiment orientation from free text customer feed-back. B. Pang did analysis of problems related to opinion mining such as opinion lexicon expansion and target extraction. Opinion targets are entities and their attributes on which opinions have been expressed. Lei Zhang et al., [10] identified domain dependent opinion words. Noun and noun phrases that indicate the

product feature which implies opinions are found using a feature based opinion mining model.

Pang et al. investigated whether binary sentiment classification can be addressed using standard topic classification techniques or not. Three classifiers, including Naive Bayes, Support Vector Machines (SVM) and Maximum Entropy are applied, to a movie review corpus by them. They also attempted to incorporate various features of the reviews into the standard bag-of- words model, such as the positions of words in the reviews, but the performance of the three classifiers was found inferior to those reported for topic classification. Pang and Lee concluded that sentiment classification is more difficult than topic classification, and that discourse analysis of reviews is necessary for more accurate sentiment analysis.

2.1 EvolvolutionOf Sentiment Analysis

Opinions Mining and sentiments analysis from natural languages is a very challenging task as it requires a deep understanding of language rules, explicit and implicit, regular and irregular behavior of languages. Sentiment analysis researchers have to face many issues with Natural Language Processing's problems. These problems cannot be resolved which may be very important. Problems like negation handling, co-reference resolution, recognition of named entity, anaphora resolution, and word-sense disambiguation.

Opinion mining is nothing but the use of natural language processing based on which the opinions can be made. Opinion mining is a NLP problem as the system mainly needs to understand the positive or negative sentiments of each and every sentence and then target the entities or topics. Therefore, sentiment analysis is a very good opportunity for NLP researchers to give concrete progress on all fronts of NLP, and potentially have a huge practical impact.

Currently, opinion mining and sentiment analysis depends on vector extraction which is used to represent the most crucial and important text features. This vector can be used to classify the relevant features which are *term frequency* and *presence*. Presence is a vector that follows binary values in which the entries indicate only whether a term occurs or not which can be given by the value 1 if it occurs and by the value 0 if it does not occur. Presence forms an effective way to review polarity classification and reveals an interesting difference. Although use of repeated keywords indicates that a topic is repeated and terms might not reflect the overall sentiment.

It is possible to add some more features like Position that refers to the token's position in a text that might affect the text's sentiment. Some methods also depend on the distance between the words. Certain textual analysis methods use part of speech information for example adjectives, adverbs, nouns, and verbs as a basic form of word-sense. Many a times adjectives works as a indicators of sentiment and guide feature selection for classification of sentiments. Some of the selected phrases chosen by already specified part of speech patterns help to detect sentiments, including an adjective or adverb.

The machine learning approach is like that of the topic of classification, with the topics being sentiment classes such as Positive and Negative. It breaks down the review statement into words or phrases, which represents the review statement as a document vector, and then classifies it based on the document vectors.

2.2Need For Sentiment Analysis

1. The increase in machine learning methods in natural language processing and information retrieval
2. The availability of datasets for machine learning algorithms which needs to be trained on, due to the tremendous advances of the World Wide Web and, especially, the development of various websites collecting information and reviews of services and products.
3. Realization of the interesting intellectual challenges & commercial and intelligence applications that the area offers.
4. Sentiment change on social media has a relation with changes in the stock market.
5. Sentiment analysis was used by The Obama team to decide public opinion to policy declaration and campaigning message ahead of his 2012 presidential election.

III.CONCLUSION

Opinion Mining is a field where the person's views on the various post can be analyzed. This analysis will be required to know the sentiments of the user and their perspective. The overall outlook towards any post which can be negative or positive can be determined. This paper gives the various ways given by authors to determine the opinion and sentiment on social media and the need of doing such analysis.

REFERENCES

1. O Harfoushi, "Sentiment Analysis Algorithms through Azure Machine Learning: Analysis and Comparison Modern Applied Science"; Vol. 12, No. 7; 2018 ISSN 1913-1844 E-ISSN 1913-1852 Published by Canadian Center of Science and Education

2Multi-Source Social Feedback of Online News Feeds NunoMoniz_and Luis Torgo LIAAD - INESC Tec FCUP - DCC, University of Porto January 23, 2018

3 B.Pang, L. Lee, and S. Vaithyanathan, "Thumbs Up? Sentiment Classification Using Machine Learning Techniques," *Proc. Ann. Conf.*

Empirical Methods in Natural Language Processing, Assoc. for Computational Linguistics, 2002, pp. 79–86.

4 B. Pang and L. Lee, "Seeing Stars: Exploiting Class Relationships for Sentiment Categorization with Respect to Rating Scales," *Proc. 43rd Ann. Assoc. for Computational Linguistics*, Assoc. for Computational Linguistics, 2005, pp. 115–124.

5 J.A. Balazs, J.D. Velásquez, "Opinion Mining and Information Fusion: A survey", Information Fusion 27 (2016) 95–110

6 A. Neviarouskaya, H. Prendinger, M. Ishizuka, Recognition of fine-grained emotions from text: an approach based on the compositionality principle, in: T. Nishida, L.C. Jain, C. Faucher (Eds.), Modeling Machine Emotions for Realizing Intelligence, Smart Innovation, Systems and Technologies, vol. 1, Springer, Berlin, Heidelberg, 2010, pp. 179–207, http://dx.doi.org/10.1007/978-3-642-12604-8_9.

7 B. Liu, Sentiment analysis and subjectivity, in: R. Dale, H. Moisl, H. Somers (Eds.), Handbook of Natural Language Processing, second ed., Machine Learning and Pattern Recognition, CRC Press, New York, NY, USA, 2010, pp. 627–666.

8 "Extracting implicit features in online customer reviews for opinion mining", Y Zhang, W Zhu - Proceedings of the 22nd International Conference on world wide web, 2013 - dl.acm.org

9 Sentiment analysis: Capturing favorability using natural language processing, T Nasukawa, J Yi - Proceedings of the 2nd *International Conference* on Data Mining, p.427, November 19-22, *2003*

10 A survey of opinion mining and sentiment analysis, B Liu, L Zhang - Mining text data, 2012 - Springer