



ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

Review on Sentiment Analysis of Social Networking Data

Aashna Rukhsaar, Dr. Anjali B.Raut

Master of Engineering, Dept. of C.S.E., HVPM's College of Engineering & Technology, Amravati. Maharashtra, India

Head of Dept, Dept. of C.S.E., HVPM's College of Engineering & Technology, Amravati. Maharashtra, India

ABSTRACT: An important part of the information era has been to seek the opinions and views of other individual. In the generation where there were no internet resources, it was customary for an individual to ask his or her friends and relatives for their opinion before making decision. Organizations conducted opinion polls, surveys to understand the sentiment and opinion of general public towards its product or services. It is necessary that sentiment analysis for customers' review division of tweets has been done to evaluate the opinion in the form of the number of tweets where opinions are highly unstructured and are either positive or negative. For this first pre-processing of dataset is done , next extracted the adjective from the dataset that have some meaning, then selected the feature vector list and thereafter applied machine learning based classification algorithms which extracts synonyms and nearer adjective for the content feature. Finally measured the performance of classifier in terms of recall (positive and negative) or under categories very bad, bad, neutral, good and very good , precision and accuracy. The sentiment analysis may also be done in human express emotions via Natural Language processing. A lot of work in Opinion Mining exists for English language. In the last few years, web contents are increasing in other languages also at a faster rate and hence there is a requirement to execute opinion mining in other languages. A Hindi Opinion Mining System (HOMS) is proposed for twitter data.

KEYWORDS: Machine Learning, Sentiment Analysis, Twitter, Public Sentiment, Hindi Language, Natural Language Processing.

I. INTRODUCTION

An important part of the information era has been to seek the opinions and views of other individual. In the generation where there were no internet resources, it was customary for an individual to ask his or her friends and relatives for their thoughts before making decision. Organizations conducted opinion polls, surveys to understand the sentiment and opinion of general public towards its product or services. In the past some years, web documents are receiving great attention as a new communication mean that describes individual thoughts, experiences and opinions. There have been a large number of research studies and industrial applications in the area of public sentiment tracking and modeling sentiment. Sentiment Analysis is the method of determining whether or not a section of writing is positive, negative or neutral. It is conjointly referred to as opinion mining, derivation the opinion perspective of a speaker. A typical case for this technology is to get however individuals feel a few particular topics. Sentiment analysis is a method where the dataset consists of emotions, attitude or assessment that takes into consideration the way an individual's thinks. The features used to categorize the sentences should have a realer strong adjective so as to summarize the review. These contents are even written in different approaches which are not easily inferred by the users or the organizations making it difficult to classify them. Sentiment Analysis influences to classify whether the information about the product is satisfactory or not before they get it. Marketers and organizations use this analysis to understand about their products or services in such a way that it can be offered as per the user's needs. Sentiment Analysis is more than just a feature in social analytics tool- it is a field of study. First, it is natural to take opinions of others before coming to a conclusion. Before the popularity of the World Wide Web, companies used to conduct polls, surveys, interviews to get the public opinion and created policies based on that. Twitter has emerged as a popular microblogging website and is being used by users to share, communicate, connect and advertise. Companies and organizations are now using this twitter data to gauge public's minds to assess the performance of their products and also to assess the general sentiment. The need is to automate and build systems that can collect, analyze and assess



ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirce.com

Vol. 4, Issue 12, December 2016

these opinions. This leads us to area of Sentiment Analysis. Sentiment Analysis or opinion mining sometimes called as one of the many areas of computational studies that deal with opinion oriented Natural Language Processing methods. Such studies include among others, genre differentiation, emotion and mood identification, ranking, relevance computations, perspective in the text source identification and opinions oriented summarization. Expressions of any sentiment or opinion are nearly defined using Natural Language Processing methods as either positive or negative. But still there are complexities. Sentiments involve emotional disposition towards particular entity. Emotions are short and brief while sentiments about a project are persistent. Thus, labeling brief emotions as enduring sentiments can be misinterpretation of it. From sentiment and opinion mining perspective, usually there are two kinds of textual information facts and opinions. Facts refer to the objective statements about the nature of a product, while opinion describes attitude, appraisals and emotions extraction of a product, service, topic or an issue. However, none of these studies performed further analysis to mine useful insights behind significant sentiment variation, called public sentiment variation. One valuable analysis is to find possible reasons behind sentiment variation, which can provide important decision-making information.

II. LITERATURE REVIEW

A. SENTIMENT ANALYSIS OF TWITTER DATA USING MACHINE LEARNING APPROACHES AND SEMANTIC ANALYSIS.

This paper proposed a set of techniques of machine learning with semantic analysis for classifying the sentence and product reviews based on twitter data[2]. Aim is to analyze a large amount of reviews by using dataset of twitter which are already labeled. The naïve byes technique which gives us a better result than the maximum entropy and SVM is being subjected to unigram model which gives a better result. Further the accuracy is improved when the semantic analysis WordNet is followed by the procedure taking it to 89.9% from 88.2%. The training data set can be increased to improve the feature vector related sentence identification process and can also extend WordNet for the summarization of the reviews. It may give better visualization of the content in better manner that will be useful for the users.

B. HOMS: HINDI OPINION MINING SYSTEM.

In this paper an Opinion mining is needed to be performed in Hindi language. Researches in the area of opinion mining and sentiment analysis have shown significant developments in the last few years[4]. These can be distinguished into two types of methods: Supervised methods and Unsupervised methods. L M Patnaik, have proposed a method to determine the opinion orientation i.e. polarity of the Hindi movie reviews.

C. WE FEEL: MAPPING EMOTION ON TWITTER.

In this paper an overview of the We Feel system for collecting and classifying emotional tweets on a global scale in real-time [5]. Over a 12-week period, 2.73×10^9 tweets were analyzed. A series of analyses were performed to demonstrate potential uses of the data, identifying the typical daily and weekly variations in emotional expression and using these to identify significant events. Correlations between emotional tweets and indices of anxiety and suicide were also observed, indicating the potential for the development of social media-based measurements for population mental to complement existing data sets. Improvements to the system and possible directions of future analysis have been suggested with the aim of better understanding the emotion of the global community and also how this relates to communities mental health and resources.

D. OPINION MINING AND SENTIMENT ANALYSIS

This paper focused on existing methods of analysis of sentiments in details i.e. supervised machine learning methods. They reduce structural risks[6]. To predict sentiment of documents, supervised machine learning approach is used here.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirce.com

Vol. 4, Issue 12, December 2016

III. PROPOSED WORK

To develop a framework that can be used to find users opinions about any product or a person by classifying the tweets into positive or negative polarity and to detect and analyze public opinions and interpreting the variations in sentiments and finding Possible reasons behind the variations in sentiments. To analyze the sentiment which are in non textual formats(Similes ,Emotics).To translate a non English word (Hindi) into English and consider it for sentiment analysis for positive or negative approach.

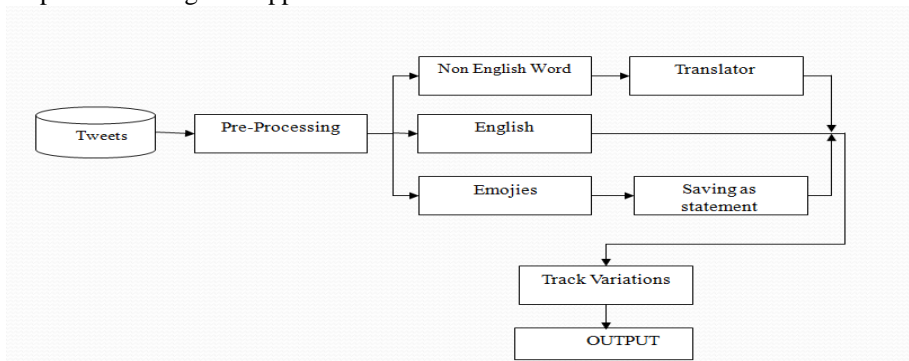


Fig 1.1.Flow of Work

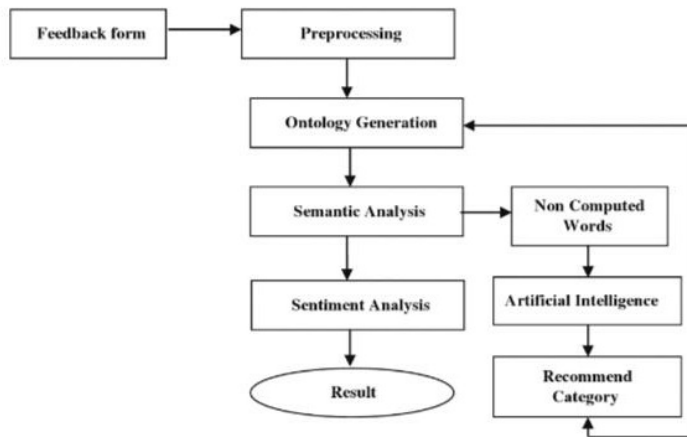


Fig1.2.System Architecture

IV. CONCLUSION

The appearance or rise in the use of micro-blogging services combined with the spread of social networking websites has established new journey with many people sharing their views and publishing their opinions on a daily basis and that too in large quantity. Sentiment Analysis of micro blogs has been widely studied in recent some years. The Previous research mainly focused on only modeling and tracking public sentiments. Our proposed models were evaluated on real Twitter data. Experimental results showed that our models can mine possible sentiment or polarity sentiment variations. Moreover, the proposed models are general: they can be used to discover special topics or aspects in one text collection in comparison with another background text collection. In addition with this the number of tweets may be observed in non English language(Hindi) , this system translated the opinions of Hindi language into English and then sentiment analysis done. Apart from this system able to performs the opinion of non textual data such as emojis into positive and negative feedback of regarding particular tweets.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 12, December 2016

REFERENCES

1. Shulong Tan ,Yang Li, Huan Sun, Ziyu Guan and Xiaofei He, "Interpreting the Public Sentiment Variations on Twitter," IEEE transaction on Knowledge and Data Engineering, Vol.26,No.5, pp.1158-1170 , 2014.
2. Geetika Gautam and Divakar Yadav,"Sentiment Analysis of Twitter Data Using Machine Learning Approaches & Semantic Analysis", IEEE conference, 2014.
3. Soujanya Poria, Erik Cambria , Alexander Gelbukh , Fedrica Bisio and Amir Hussain , "Sentiment Data Flow Analysis by Means of Dynamic Linguistic Patterns", IEEE Computational Intelligence Magazine, November 2015.
4. Vandana Jha , Manjunath N* and Deepa Shenoy, "HOMS: Hindi Opinion Mining System",IEEE international conference on (ReTIS),2015.
5. Mark E. Larsen, Tjeerd W. Boonstra, Philip J. Batterham, Bridianne O'Dea, Cecile Paris, and Helen Christensen,"We Feel: Mapping emotion on Twitter", IEEE Journal of Biomedical and Health Informatics-2014.
6. B. Pang and Lillian Lee,"Opinion Mining and Sentiment Analysis", Foundation and Trends in Information Retrieval, Vol 2,pp 1-135, 2008
7. J. Bollen, H. Mao, and A. Pepe, "Modeling public mood and emotion:Twitter sentiment and socio-economic phenomena," in Proc. 5th Int. AAI Conf. Weblogs Social Media, Barcelona, Spain, 2011.
8. J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," J. Comput. Sci., vol. 2, no. 1, pp. 1–8, Mar. 2011.
9. D. Chakrabarti and K. Punera, "Event summarization using tweets," in Proc. 5th Int. AAI Conf. Weblogs Social Media, Barcelona, Spain, 2011.
10. A. Go, R. Bhayani, and L. Huang, "Twitter sentiment classification using distant supervision," CS224N Project Rep., Stanford: 1–12, 2009.
11. T. L. Griffiths and M. Steyvers, "Finding scientific topics," in Proc. Nat. Acad. Sci. USA, vol. 101, (Suppl. 1), pp. 5228–5235, Apr. 2004.
12. D. Hall, D. Jurafsky, and C. D. Manning, "Studying the history of ideas using topic models," in Proc. Conf. EMNLP, Stroudsburg, PA, USA, 2008, pp. 363–371.
13. G. Heinrich, "Parameter estimation for text analysis," Fraunhofer IGD, Darmstadt, Germany, Univ. Leipzig, Leipzig, Germany, Tech. Rep., 2009.
14. Z. Hong, X. Mei, and D. Tao, "Dual-force metric learning for robust distracter-resistant tracker," in Proc. ECCV, Florence,Italy, 2012.
15. T.Sakaki, M.Okazaki and Y. Matsuo, "Earthquake shakes Twitter Users: Real Time Event Detection by Social Sensors", Proceedings of International Conference on World Wide Web, pp 851-860, USA, 2010.
16. Luciano Barbosa and Junlan Feng,"Robust Sentiment Detection on Twitter from Biased and noisy Data", 23rd International Conference on Computational Linguistic, pp 36- 44, 2010.
17. Long Jiang, Mo Yu, Ming Zhou, Xiaohua Liu and Tiejun Zhao, "Target dependent Twitter Sentiment Classification", Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics, pp 151–160, 2011.
18. M. Hu and B. Liu, "Mining and summarizing customer reviews,"
19. in Proc. 10th ACM SIGKDD, Washington, DC, USA, 2004.
20. Y. Hu, A. John, F. Wang, and D. D. Seligmann, "Et-lda: Joint topic modeling for aligning events and their twitter feedback," in Proc. 26th AAAI Conf. Artif. Intell., Vancouver, BC, Canada, 2012.
21. L. Jiang, M. Yu, M. Zhou, X. Liu, and T. Zhao, "Target-dependent twitter sentiment classification," in Proc. 49th HLT, Portland, OR, USA, 2011.
22. M.Hu and B. Liu, "Mining and Summarizing Customer Reviews", 4th proceeding International Conference on Knowledge Discovery and Data Mining, pp 168- 177, 2004. [22] V.P.H Binali and w.chen, "A State of The Art Opinion Mining and its Application Do mains", IEEE International Conference on Industrial technology, pp 1-6, Feb 2009.
23. V.P.H Binali and w.chen, "A State of The Art Opinion Mining and its Application Do mains", IEEE International Conference on Industrial technology, pp 1-6, Feb 2009.

BIOGRAPHY



Aashna Rukhsaar is a Master of Engineering Student in the Computer Science & Engineering Department,H.V.P.M'S College of Engineering & Technology,Amravati, Maharashtra. She received Bachelor of Engineering Degree in 2015 from SGBAU Amravati, Maharashtra,India. Her research interests are Data Mining.



Dr. Anjali B.Raut is a Head Of Department in the Computer Science & Engineering, ,H.V.P.M'S College of Engineering & Technology, Amravati, Maharashtra. She received PhD in Data Mining in 2014 from SGBAU Amravati, Maharashtra, India.