iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 2 Issue: XII Month of publication: December 2014

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

www.ijraset.com Volume 2 Issue XII, December 2014
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Interpreting the Public Sentiment with Emotions on Twitter

Ms. Waghode Poonam B[1], Prof. Mayura Kinikar[2]
[1,2] *MIT Academy of Engineering, Alandi, Pune, Maharashtra, India*

Abstract: The Information on the Web is the ocean of opinions expressed in the user generated form as tweets on twitter which explore a great opportunity to understand the sentiment of the public by analyzing its large-scale data as well as opinion-rich data. Now people increasingly use emoticons in tweets in order to express their feelings on wide variety of topics on Twitter. Sentiment analysis on entities in tweets thus becomes a fast and effective way of concluding public opinion. But the tweets containing text is not sufficient for the opinion but somewhere emoticons also place a very vital role in the field of Sentiment Analysis. Emoticons are widely used to express positive or negative sentiment on Twitter. Inspired by the wide availability of emoticons, we propose to study the literature survey for sentiment analysis on tweets with emoticons.
Keywords—Sentiment analysis, tweets, Emoticons, Unsupervised Approach

I. INTRODUCTION

With the explosive growth of user generated messages, Social site like Twitter where millions of users can exchange their opinion regarding some topic. Sentiment analysis on Twitter data has provided a platform where timely public sentiment can be expose in an economical and effective way, which is critical for decision making in various domains. For instance, a company can study the public sentiment in Tweets to obtain users' feedback towards its products; while a politician can adjust his/her position with respect to the sentiment change of the public.

Sentiment analysis is a technique for extracting sentiment associated with polarities of positivity, negativity and neutrality. It is one of the types of natural language processing in which we can track the mood of the public about a particular entity. Sentiment analysis, which is also called opinion mining, is used for constructing a system to collect and examine opinions about the entities on posts in Twitter called *tweets*. Due to the explosion of social media services a great opportunity to understand the sentiment of the public by analyzing its large-scale data as well as opinion-rich data.Sentiment analysis on tweets can done by many approaches. Approaches as machine-learning and lexicon-based approaches have been adopted to do sentiment analysis on Twitter. Machine-learning approaches to sentiment analysis need to train the data
and are known to be domain-dependant. Lexicon-based sentiment analysis approaches use sentiment lexicons for retrieving the polarity of individual words and aggregate these scores in order to determine the text's polarity. But the tweets containing text is not sufficient for the opinion but somewhere emoticons also place a very vital role in the field of Sentiment Analysis Emoticons are widely used to express positive or negative sentiment on Twitter .We propose to study the literature survey on unsupervised approach for sentiment analysis with emoticons.

The remainder of the paper is organized as follows. Section 2 illustrates Emoticons. Section 3 provides a brief review of Literature Survey. Sentiment Analysis Emoticons Framework is reported in Section 4.

II. EMOTICONS

Sentiment analysis on entities in tweets thus becomes a rapid and effective way of concluding public opinion. But the tweets containing text is not sufficient for the opinion but somewhere emoticons also place a very vital role in the field of Sentiment Analysis. Emoticons are widely used to express positive or negative sentiment on Twitter. The first emoticon was used on September 19, 1982 by professor Scott Fahlman in a message on the computer science bulletin board of Carnegie Mellon University. In his message, Fahlman proposed to use ":-)" and ":-(" to distinguish jokes from more serious matters, respectively. It did not take long before the phenomenon of emoticons had spread to a much larger community.

Sentence	How	Sentiment
I love my work :-D	Intensification	Positive
The movie was bad :-D	Negation	Positive
:-D I got a promotion	Only sentiment	Positive
-_- I love my work	Negation	Negative
The movie was bad -_-	Intensification	Negative
I got a promotion -_-	Only sentiment	Negative

Table 1: Example of how Emoticons can be used to convey sentiment

ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

People started sending yells, hugs, and kisses by using graphical symbols formed by characters found on a typical keyboard. A decade later, emoticons had found their way into everyday computer-mediated communication and had become the paralanguage of the Web [17]. As such, emoticons enable people to indicate subtle mood changes, to signal irony, sarcasm, and jokes, and to express, stress, or disambiguate their (intended) sentiment, perhaps even more than nonverbal cues in face-to-face communication can. Therefore, harvesting information from emoticons appears to be a viable strategy to improve the state-of-the-art of sentiment analysis.

III. LITERATURE SURVEY

Sentiment analysis has been a hot topic for quite a few years [2]. This paper clarifies the differences between affect, feelings, emotions, sentiments, and opinions these five subjective terms and reveals significant concepts to the computational linguistics community for their effective detection and processing in text[1].

Author perceive that guiniune reasons behind variations is highly related to the emerging topics within the sentiment variation periods. Based on this observation, author propose a Latent Dirichlet Allocation which is also called as LDA based model, Foreground and Background LDA which is called as FB-LDA, to extract foreground topics and filter out old toics which stayed for long time background topics[2].[3]This paper presents an interactive visualization system that aims to analyze public sentiments for popular topics on the Internet. Author illustrated that by searching and to establish frequent words in text type data, it models and mines the changes of the sentiment on topics[4]. Authors propose a method for automatic analysis of attitude (affect, judgment, and appreciation) in sentiment words. They proposed first stage as an automatic separation of clear affective and relating to (judgmental) adjectives from those that express appreciation or different attitudes depending on context[5].

Recently, as an effective tool to understand opinions of the public, sentiment analysis is widely used in various social media applications [7], including poll rating prediction [4], stock market prediction, event analysis [1], behavioral targeting [3], etc. Similar to conventional sentiment analysis on product and movie reviews, most existing methods in social media can fall into supervised learning methods [10] and unsupervised learning methods [4].Due to the lack of label information and the large-scale data produced by social media services, unsupervised learning becomes more and more important in real-world social media applications. As in many existing system they are depended on one single domain only using supervised technique .But as unsupervised technique is also efficient to use in any domain without need of any training data.

In existing research [1] author have described the lexicon based classifier with some future work regarding Emoticons, Natural Language Processing.[2] In this paper author has proposes a Model for general people's opinions in regard to Australian federal election 2010 event was taken as an example for sentiment analysis experiments.[13] In this paper ,author present an unsupervised system for extracting aspects and determining sentiment in review text. Author has demonstrate its effectiveness on both component tasks, where it achieves similar results to more complex semi-supervised methods that are restricted by their reliance on manual annotation and extensive knowledge sources.[19]In this paper, author have presented a method for an automatic collection of a corpus that can be used to train a sentiment classifier. They used TreeTagger for POS-tagging and observed the difference in distributions among positive, negative and neutral sets. From the observations it is concluded that authors use syntactic structures to describe emotions or state facts. The limitations as it only reliable for Chinese language.

IV. SENTIMENT ANALYSIS EMOTICONS FRAMEWORK

We propose a novel framework for automated sentiment analysis, which takes into account the information conveyed by emoticons. The goal of this framework is to detect emoticons, determine their sentiment, and assign the associated sentiment to the selected text in order to correctly classify the polarity of natural language text as either positive or negative. In order to accomplish this, we build upon existing work [2]. Our framework, depicted in Figure 1, is essentially a pipeline in which each component fulfills a specific task in analyzing the sentiment of tweets.

First, we load a document in order for it to be analyzed for sentiment. we look for empty lines or lines starting with an indentation. When splitting a document into sentences, we look for punctuation marks, such as \.", \!", and \?", as well as for emoticons, as most emoticons are placed at the end of a text segment. Each text segment is checked for the presence of emoticons (step 2). If a word in a text segment matches a character sequence in the emoticon sentiment lexicon, the segment is rated for sentiment based on the sentiment imposed onto the text by its emoticons. Else, the segment is analyzed for the sentiment conveyed by its sentiment-carrying words that id pre-processing, Feaure Selection and Text based Sentiment Analysis.

Given a set of tweets, T , which contains a set of sentences, s, T = {s1 , s2 , … , si}; and each sentence sk describes something on a subset of entities e = {ei , ..., ej|ei , ej Ɛ E}, where E is the set of all entities. set of opinion word, w, s

= {w1, w2,… ,wj}. The score function for a sentence is as follow:

$$score(s) = \sum_{w_j : w_j \in s \wedge w_j \in WL} \frac{w_j \cdot sentOri}{dis(w_j, e_i)} \quad (1)$$

(1)

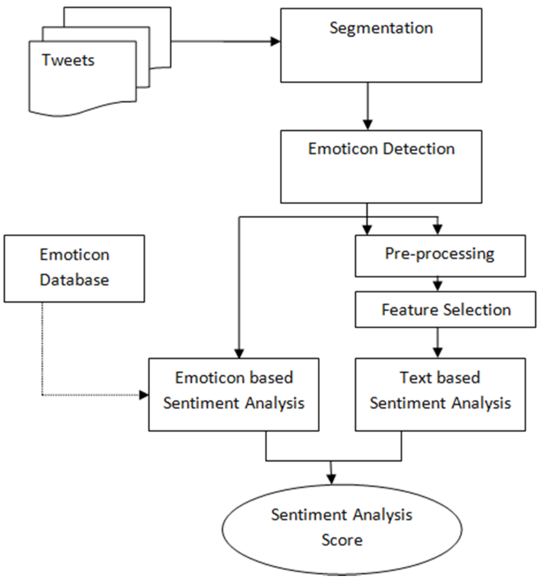

Figure 1: Framework of Model

where wj is an opinion word, WL is the set of all opinion words from Wilson lexicon list and s is the sentence that contains the entity ei , and dis (wj , ei) is the distance between entity ei and opinion word wj in the sentence s, and wj.sentOri is the semantic orientation of the word wj (i.e.,+1, or +0.5, or 0, or −1, or −0.5). If a sentence contains more than one entity then the opinion word close to the entity has smaller value of dis (wj , ei) and indicates this word makes more contribution to that entity's sentiment scores.

The Mathematical Model of the system will be given as-
Score={wj,ei,WL,SentOri,s}where,

- wj=Opinion word
- ei=Entity
- SentiOri=Semantic Orientation
- WL=Wordnet List
- s=Sentence that contain an entity

V. ACKNOWLEDGEMENT

I thank Mrs. Mayura Kinikar for guiding and helping for the project information and valuable comments. She helped in a broad range of issues from giving me direction, helping to find the solutions, outlining the requirements and always having the time to see me.
I have furthermore to thank Prof. R. M. Goudar, M.E. Coordinator Department of Computer Engineering, to encourage me to go ahead and for continuous guidance. I also want to thank Prof. Uma Nagraj for all her assistance and valuable guidance.

REFERENCES

[1] Myriam Munezero, Calkin Suero Montero,Erkki Sutinen, and John Pajunen, "*Are They Different? Affect, Feeling, Emotion,Sentiment, and Opinion Detection in Text*", IEEE transactions on affective computing, pp. 101-111 VOL. 5, NO. 2, APRIL-JUNE 2014

[2] Shulong Tan, Yang Li, Huan Sun, Ziyu Guan, Xifeng Yan, Jiajun Bu,Chun Chen,and Xiaofei He, "Interpreting the Public Sentiment Variations on Twitter", IEEE transactions on knowledge and data engineering,pp. 1158-1171, VOL. 26, NO. 5, MAY 2014

[3] Georgios Paltoglou, Mathias Theunis, Arvid Kappas, and Mike Thelwall, "Predicting Emotional Responses to Long Informal Text", IEEE transactions on affective computing,pp.106-116 vol. 4, no. 1, january-march 2013

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

[4] Changbo Wang, Zhao Xiao, Yuhua Liu, Yanru Xu, Aoying Zhou, and Kang Zhang, "SentiView: Sentiment Analysis and Visualization for Internet Popular Topics" ,IEEE transactions on human-machine systems,pp.620-631,vol. 43, no. 6, november 2013

[5] Alena Neviarouskaya and Masaki Aono, "Sentiment Word Relations with Affect,Judgment, and Appreciation",IEEE transactions on affective computing,pp. 425-439 vol. 4, no. 4, october-december 2013

[6] Alexander Hogenboom, Daniella Bal, Flavius rasincar, Malissa Bal, Franciska de Jong,Uzay Kaymak, *"Exploiting emoticons in sentiment analysis*". In Proceedings of the 28th Annual ACM Symposium on Applied Computing, Pp. 703-710, 2013

[7] Weichselbraun, A. ; Gindl, S. ; Scharl, "Extracting and Grounding Contextualized Sentiment Lexicons",In Intelligent Systems, IEEE Trans.,Vol.28, No. 2, pp: 39 - 46,March-April 2013

[8] Georgios Paltoglou and Mike Thelwall. "*Twitter, MySpace,D igg: Unsupervised Sentiment Analysis in Social Media"* ACM Transactions on Intelligent Systems and Technology, Vol. 3, No. 4, Article 66, Publication date: September 2012.

[5] Xujuan Zhou,Xiaohui Tao,Jianming Yong and Zhenyu Yang. "*Sentiment Analysis on Tweets for Social Events*". In Proceedings of the IEEE 17th International Conference on Computer Supported Cooperative Work in Design ,pp. 557-562,2013.

[6] S. Brody and N. Elhadad. "*An unsupervised aspect- sentiment model for online reviews*". In HLT '10: Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics, pp. 804-812, Morristown, NJ, USA, 2010. Association for Computational Linguistics.

[7] K. Dave, S. Lawrence, and D. Pennock. "*Opinion extraction and semantic classification of product reviews*". In Proceedings of the 12th International World Wide Web Conference (WWW), pp. 519-528, 2003

[8] E.Schonfeld. Mining the thought stream. TechCrunch Weblog Article, *http://techcrunch.com/2009/02/15/mining-the-thought- stream//, 2009.*

[9] V. Hatzivassiloglou and J. M.Wiebe. "*Effects of adjective orientation and gradability on sentence subjectivity*". pp. 299-305, 2000.

[10] B. J. Jansen, M. Zhang, K. Sobel, and A. Chowdury. "*Twitter power: Tweets as electronic word of mouth*". J. Am. Soc. Inf. Sci., 60(11):2169-2188, 2009.

[11] M. M. Bradley and P. J. Lang. "*Affective norms for english words (anew): Stimuli,instruction manual, and affective rating*"s. Technical report, Center for Research in Psychophysiology, University of Florida, Gainesville, Florida, 1999.

[12] J. Kamps, M. Marx, R. Mokken, and M. de Rijke. "*UsingWordNet to measure semantic orientation of adjectives*". In Proceedings of the 4th International Conference on Language Resources and Evaluation, pages 1115-1118, 2004.

[14] B. Liu. "*Handbook of Natural Language Processing, chapter Sentiment Analysis and Subjectivity*". Second edition, 2010.

[15] S. Milstein, A. Chowdhury, G. Hochmuth, B. Lorica, and R. Magoulas. "*Twitter and the micro-messaging revolution*": Communication, connections, and immediacy140 characters at a time. An OReilly Radar Report. 54 pp., November 2008.

[16] B. Pang and L. Lee. "*Thumbs up? sentiment classification using machine learning techniques*". In Proceedings of EMNLP, pp. 79-86, 2002.

[17] B. Pang and L. Lee. "*Opinion mining and sentiment analysis*". Foundations and Trends in Information Retrieval, pp. 1-135, Jan. 2008.

[18] A.-M. Popescu and O. Etzioni. "*Extracting product features and opinions from reviews*". In HLT '05:Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing, pp. 339 - 346, Morristown, NJ, USA, 2005.

[19] P. D. Turney and M. L. Littman. "*Unsupervised learningof semantic orientation from a hundred-billion-word corpus*". Dec 2002.

[20] J.Wiebe. "*Learning subjective adjectives from corpora*".In Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence, pp.735-740. AAAI Press, 2000.

[21] D. T. Wijaya and S. Bressan. *"A random walk on the red carpet: rating movies with user reviews and page rank".* In CIKM '08: Proceeding of the 17th ACM conference on Information and knowledge management, pp. 951-960. ACM, 2008.

Waghode Poonam B has received B.E[Computer Engineering] from North Maharashtra University in 2009 & pursuing M.E[Computer Engineering] From M.I.T, AOE, Alandi. Her research interest includes Data Mining

Mayura.U.Kinikar has received M.E [Comp Science & Engineering] from Government College of Engineering Aurangabad. Her research interest includes Data Mining.She is working as Asst. Prof. at M.I.T, AOE, Alandi, Pune.