Sentiment Analysis of Indian Government Schemes Using Twitter Datasets

Bharat R.Naiknaware, Seema Kawathekar, Sachin N.Deshmukh

Abstract: *The Social media monitoring has been growing day by day so analyzing of social data plays an important role in knowing peoples behavior. So we are analyzing Social data such as Twitter Tweets using sentiment analysis which checks the opinion of peoples related government schemes that are announced by Central Government. This paper develops based on social media twitter datasets of particular schemes and its polarity of sentiments. Popularity of Internet has been rapidly increased. Sentiment analysis and opinion mining is the field of study that analyses people's opinions, sentiments, evaluations, attitudes, and emotions from written language. User generated contents are highly generated by users. The growing importance of sentiment analysis coincides with the growth of social media such as reviews, forum discussions, blogs, micro-blogs, Twitter, and social networks. It is difficult to analysis or summarize the user generated content. Most of the users writes their opinions, thoughts on blogs, social media sites, E-commerce site etc. So These Contents are very important for individuals, industry, Government and research work to take decisions. For this Sentiment analysis and opinion mining research is hot research area which comes under the Natural Language processing. We plot and calculate numbers of positive, negative and neutral tweets from each events.*
Keywords: *Sentiment Analysis and Opinion mining, Natural language processing.*

I. Introduction

Social media have become an emerging phenomenon due to the huge and rapid advances in information technology. People are using social media on daily basis to communicate their opinions with each other about wide variety of subjects, products and services, which has made it a rich resource for text mining and sentiment analysis. Social media communications include Facebook, twitter, and many others. Twitter is one of the most widely used social media sites.

The number of twitter messages sent per second worldwide. There is no standard method for mining and analyzing social media business data. Here, an open source approach for text mining and sentiment analysis using a set of R packages for mining twitter data and sentiment analysis is presented, which is applicable for other social media sites. A case study of Eight Indian Government Schemes is presented to show the importance of analyzing user generated online opinions from Twitter. This is helpful for evaluation government performance monitoring from Peoples perspective instead of making Peoples surveys which are expensive and time consuming. Sentiment analysis has been first introduced by Liu, B. It is also known as opinion mining and subjectivity analysis is the process to determine the attitude or polarity of opinions given by humans to particular scheme. Sentiment analysis can be applied on any textual form of opinions such as blogs, reviews and Microblogs. Microblogs are those small text messages such as tweets, a short message that cannot exceed 160 characters. These microblogs are easier than other forms of opinions for sentiment .Sentiment analysis can be done on a document level or a sentence level. In the first case, the whole document is evaluated to determine the opinion polarity, where, the features describing the product/service should be extracted first. Whereas, the second one, the document is divided into sentences each one is evaluated separately to determine the opinion polarity.

1.1 Levels of Sentiment Analysis
a. Document level Sentiment Analysis

In this Sentiment Analysis level whole document has analyzed and classify whether a whole opinion document expresses a positive or negative sentiment [1], [2]. In one document only reviews of one product has been reviewed. And task is to find out the opinion about that product. So this task is broadly known as document-level sentiment classification. In this level, expressed opinion is on single entity. This is not applicable when there is document which contains multiple product reviews.

b. Sentence Level Sentiment Analysis

In this level, task goes to every sentence and determine whether the sentence expresses the positive, negative or neutral opinion. This level attentively related to Subjectivity Classification [3], which distinguishes objective Sentences and subjective sentences. Objectives sentences express factual information about sentences where

Subjective sentences express the subjective information about sentences. Many objective sentence can involve Opinions. This task is known as Sentence Level Sentiment Analysis.

c. Aspect level Sentiment Analysis

Aspect Level sentiment Analysis was earlier called Feature level (feature-based opinion mining and summarization) Sentiment Analysis [4]. Document and Sentence Level Sentiment Analysis do not find out what exactly people like or did not like. It achieves finer-grained analysis. In this level directly looks at the opinion itself instead of looking to documents, paragraphs, sentences, clauses or phrases. This level consider the entity, aspect of that entity, opinion of aspect, opinion holder and time. Because of these parameters this level can find what actually people like means which feature of product mostly likes by customers and also on which time. This task is more interesting and more difficult too.

II. Data Source

Implementations of different government schemes and quality of government schemes is depends upon the people's opinion .Using this opinions individual or government can find out the popular Government Schemes and feedback of peoples from Internet. User are giving their opinions on different social media websites like twitter, Facebook, LinkedIn etc. Many of these blogs contains different reviews of different Schemes, events, Amendments etc.

a. Blogs

Blog is a web site on which someone writes about personal opinions, activities, and experiences. Blogging is one of the most valuable tools that government have to communicate with peoples and ultimately make their lives easier. With the increasing growth of User generated Content on the internet, blogging pages are increasing rapidly. For expressing personal opinion Blog pages are most popular. Bloggers record the daily events in their lives and express their opinions, feelings, and emotions in a blog [5]. In Sentiment analysis and opinion mining research blogs are used as source of people's opinion and used in implantation of schemes.

b. Data Set

We collect data from twitter microblogs and also find the related to government scheme data which is available on website (https://data.gov.in/policies).

c. Review sites

Opinion can be an important factor in making a decision for any peoples. A large and unstructured data is a growing body of user-generated reviews are available on the Internet. Government schemes popularity are based on opinions expressed in much unstructured format. People's data for research is mostly crawled from different social Medias such as websites, blogs, discussion forums etc.

d. Micro-blogging

Twitter is a popular microblogging service where users create status messages called "tweets". These tweets contains people's opinion, thoughts on different schemes. Tweets are sometime used in classifying the sentiment about particular scheme

III. Experimental Setup

Concern research work it's related to Sentence level sentiment analysis that on Government Schemes are announced by Central government's authority. There are different machine learning techniques for sentiment Analysis such as Naive Bayes, maximum entropy classification, and support vector machines etc. for this research work here applied Naïve Bayes techniques and dataset collected from social site i.e. Twitter. We have taken datasets concern with eight government schemes from twitter and we analyze that and observed the outcome experimental results are quite good in comparison to the human generated datasets. SVM works best as compare to other classifiers. In [2] present a simple unsupervised learning algorithm for classifying a tweets as recommended or not recommended. Average Semantic orientation of the phrases in the tweets that contain adjectives or adverbs is used for classifying the tweets. This paper works on Sentence level Sentiment Analysis. Point wise Mutual Information (PMI) is used calculate semantic orientation of phrase and the word. We took tweets from Opinions for different Schemes Algorithm designed for isolated adjectives, rather than phrases containing adjectives or adverbs. They used four step supervised learning algorithm to infer the semantic orientation of adjectives from constraints on conjunctions. In that they got accuracy for classification of adjectives it tracks online tweets on Schemes and generate plot which contains number of positive sentiment and negative sentiment messages over time. Schemes as either positive or negative to predicting based on polarity of datasets. They checked human performance at the task. Applied algorithm is Meta algorithm, Based on a metric labelling. This Meta algorithm can give best performance over both multi-class and regression versions of

SVMs when we employ a novel similarity measure appropriate to the problem. Determined the opinions or sentiments expressed on different schemes, e.g., of a #Digital India, #GST4India etc. They analyses the different aspect of different schemes. This paper performed three tasks i.e. (1) mining schemes features that have been commented on by peoples (2) identifying opinion sentences in each scheme and deciding whether each opinion sentence is positive or negative (3) summarizing the results. In their proposed technique we did POS tagging, Frequent Features Identification, Opinion Words Extraction, and Orientation Identification for Opinion Words, Infrequent Feature Identification, Predicting the Orientations of Opinion Sentences and Summary Generation. They got the results of opinion sentence extraction and sentence orientation prediction.

3.1 Sentiment Analysis Process
Sentiment Analysis is classified into two main approaches i.e. Supervised Learning Approach and Unsupervised Approach. In Sentiment Analysis Process Following Steps are necessary [1].

a. Collection of Peoples tweets
Tweets are necessary for doing the Sentiment Analysis Task. For the Collection of tweets there are different techniques which are used in this survey. Schemes tweets are collected from Twitter websites. The tweets can be a structured, semi-structured and unstructured type. Sentiment Analysis research, there are open source framework where researcher can get their data for the research purpose. R [27] is one of the programming language and software environment for statistical computing and graphics supported by the R Foundation for Statistical Computing. By installing required packages and authentication process of social website, to crawl the reviews from that site is easy task. Once we have our text data with us then we can use that data for Pre-processing purpose.

b. Pre-Processing
Data pre-processing is done to remove the incomplete noisy and inconsistent data. Data must be preprocessed before using in feature selection task. In pre-processing following are some tasks:
• Removing URLs, Special characters, Numbers, Punctuations etc.
• Removing Stopwords
• Removal of Retweets (in case of twitter dataset)
• Stemming
• Tokenization

c. Feature Selection
Feature selection from pre-processed text is the difficult task in sentiment analysis. The main goal of the feature selection is to decrease the dimensionality of the feature space and thus computational cost. Feature selection will reduce the overfitting of the learning scheme to the training data. In [1] different machine learning algorithms were analyzed on a Scheme dataset with different feature selection techniques features.

d. Sentiment Word Identification
Sentiment word identification is a fundamental work in numerous applications of sentiment analysis and opinion mining, such as tweets mining, opinion holder finding, and tweet classification. Sentiment words can be classified into positive, negative and neutral words.

e. Sentiment Polarity Identification
The basic task in Sentiment Analysis is classifying the polarity of a given tweets feature. The polarity is in three category i.e. Positive, Negative and Neutral. Polarity identification is done by using different lexicons e.g. Bing Lui sentiment lexicon, SentiWordNet etc. which help to calculate sentiment score, sentiment strength etc.

f. Sentiment Classification
Sentiment classification of government schemes tweets dataset and opinion of schemes dataset is done using supervised machine learning approaches like naïve Bayes, SVM, Maximum Entropy etc. Accuracy is depends on which dataset is used for which classification methods. In the case of Supervised machine learning approaches Training dataset is used to train the classification model which then help to classify the test data.

g. Analysis of Reviews
Finally Analysis of result is important to make decision to individual and government. In case of government schemes that are announced by central government tweets if more result is positive then peoples can like to this particular schemes Analysis is used in taken feedback of particular schemes that are take appropriate discussion to public and take decision for proper implementation of government schemes
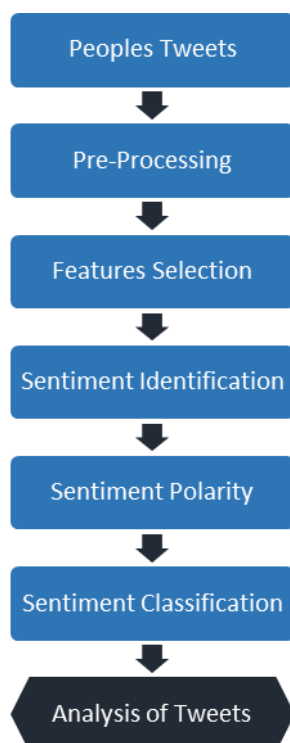
Figure 1: Sentiment Analysis Preprocessing Approach
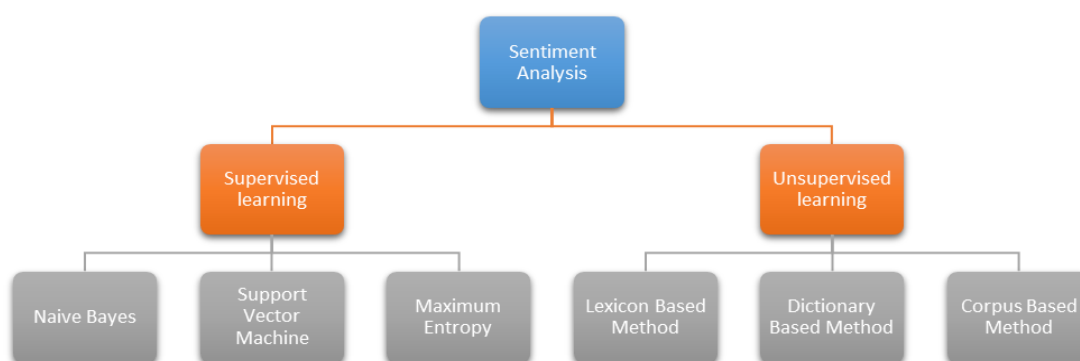
3.1 Sentiment Classification Approaches [1]

Figure 2: Sentiment Classification Approaches There are two main approaches in sentiment analysis i.e. Supervised learning and Unsupervised Learning Approach

3.2.1 Unsupervised Learning Approach

This approach is used to classify the sentiment when we have training data and it can solve the problem of domain dependency and need to reduce the training data. Turney [2] used two seed words (poor and excellent) to calculate the semantic orientation of the phrase. Point wise mutual information has been used to find out association of seed words with their phrase. Sentiment of document is calculated as the average semantic orientation of all such phrases. This approach got 66% accuracy for the movie review dataset at document level.

A) Lexicon based Methods

Lexicon-based approaches for sentiment classification are based on the insight that the polarity of a piece of text can be obtained on the ground of the polarity of the words which compose it. [28] In this methods lexical resources which are concerned with mapping words to a categorical (Positive, Negative, Neutral) or numeric score calculated by the algorithm to get the overall sentiment of that text. Following are some lexical resources:

a. SentiWordNet

SentiWordNet [29] is a lexical resource used in the sentiment analysis applications. It gives an annotation based on three numerical sentiment scores (positivity, negativity, neutrality) for each WordNet synset [30]. It provides a synset-based sentiment representation, different senses of the same Term may have different sentiment scores. Word Sense Disambiguation (WSD) algorithm to identify the most promising meaning.

b. WordNet-Affect

WordNet-Affect [16] is a linguistic resource for a lexical representation of affective knowledge. It is an extension of WordNet which labels affective-related synsets with affective concepts defined as A-Labels (e.g.the term euphoria is labelled with the concept positive-emotion, the noun illness is labelled with physical state, and so on). The mapping is performed on the ground of a domain-independent hierarchy.

c. MPQA

It [18] is related to subjectivity lexicons and provides lexicons of 8,222 terms labelled as subjective expressions which are gathered from different sources. This contains list of words along with their POS-tagging, labelled with polarity (positive, negative, neutral) and intensity (strong, weak).

d. SenticNet

This [19] lexical resource is used for concept-level sentiment analysis. It depends on Sentic Computing [34] which is multi-disciplinary paradigm for Sentiment Analysis. SenticNet is able to associate polarity and affective information also to complex concepts like accomplishing goal, celebrate special occasion and so on. It gives sentiment score in range between -1 and 1 for 14,000 common sense concepts.

e. Hu and Liu's lexicon

This opinion lexicon contains a list of positive and negative words or sentiment words for English. This list was compiled for [4] paper which contains 2006 positive and 4783 negative sentiment words.

B) Dictionary Based Methods

f. WordNet

WordNet [22] is a large lexical database of English. Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. Synsets are interlinked by means of conceptual-semantic and lexical relations. WordNet superficially resembles a thesaurus, in that it groups words together based on their meanings. In [36] polarity of a word is determined by measuring its shortest distance to „good" and „bad". We extract the words that are contained by WordNet from our dictionary for comparison in our experiment.

g. General Inquirer

The General Inquirer (GI) is an application in text analysis with one of the oldest manually constructed lexicons. The GI has been in development and refinement since 1966, and is designed as a tool for content analysis, a technique used by social scientists, political scientists, and psychologists for objectively identifying specified characteristics of messages [26]. The lexicon contains more than 11K words classified into one or more of 183 categories. In GI there are 1,915 words labelled Positive and the 2,291 words labelled as Negative. It is ideally used in several research to automatically determine sentiment properties of textual data.

h. LIWC

LIWC is [29] text analysis software designed for studying the various emotional, cognitive, structural, and process components present in text samples. LIWC uses a proprietary dictionary of almost 4,500 words organized into one (or more) of 76 categories, including 905 words in two categories especially related to sentiment analysis. First category is Positive Emotions (e.g. Love, nice, good, great) which are 406 in numbers and Negative Emotions (Hurt, ugly, sad, bad, worse) which are 499 in number.

i. AFINN

AFINN [20] is a list of English words rated for valence with an integer between minus five (negative) and plus five (positive). The words have been manually labelled by Finn Årup Nielsen in 2009-2011. The file is tab separated. In first version AFINN-96 1468 unique words and phrases on 1480 lines and in second updated version AFINN-111 2477 words and phrases are present.

C) Corpus Based Methods

j. Darmstadt Service Review Corpus

It consists of consumer reviews annotated with opinion related information at the sentence and expression levels. In [23] they gave Niek Sanders: He has constructed a Twitter Sentiment Corpus that "consists of 5513 hand-classified tweets.

IV. Proposed Methodology

The methodology used for mining twitter dataset is shown in fig 3 and following steps are important in this Methodology.

a. Data Collection

R is a programming language and software environment for statistical computing and graphics supported by the R Foundation for Statistical Computing [20]. R has different packages which helps to get the social media data like Twitter, Facebook and also different packages for pre-processing on text / numeric data and visualization of results i.e. tm, stringr, ggplot, NLP, Curl etc. TwitteR package is used to get tweets from twitter using twitter API.

b. Pre-Processing or Data Cleaning

Pre-processing is very important in data mining and it effects on the accuracy of result. In Preprocessing tm package is used to get text from tweets and other pre-processing steps like Stopwords removal, removing spaces, punctuation, URLs and performing stemming (get the root of the words). After this step unstructured data represents in Term-Document Matrix.

c. Data Analysis

From the above step we got TDM and using TDM, it is easy to find association rules, finding more frequent terms and performing sentiment analysis using the lexicon-based approach, which uses a set of positive and negative words. Using Scoring Function score of every tweet has been calculated using Bing Lui lexicons [4].

V. Experiments And Result

The experiments having collection of different hashtag (#) tweets. TwitteR package used to access the live tweets from twitter. ROAuth, TwitteR, Rcurl etc. packages enables authentication and access to twitter messages by using keyword search queries [27].

Table 1. Twitter Dataset Used for Experimental Work

| Trends	No of Tweets	Year
#GST4India	5000	2016
#startupindia	2500	2016
#SwachBharat	5000	2016
#FinBudget	25000	2016
#Digital india	5000	2016
#RailwayBudget	4000	2016
#MakeInIndia	10000	2016
#Kashmir	4000	2016

It used Unsupervised Learning approaches which contains lexicons in this methodology. Tweet has been accessed over time with different topics. Related to sentiment analysis and opinion mining for in this methodology lexicon based methods has been used. Fig 1 shows the Wordcloud of #MakeInIndia topic. For calculating score of each tweet we need the opinion lexicons and scoring function. Working of scoring function is as follows:

Sentiment Score = □ positive words– □ Negative words

Score will be positive if the number of positive words are greater than number of negative words which called as positive polarity. Score will be negative if the number of negative words are greater than number of positive words which called as negative polarity. Score will be neutral if the number of positive and negative words are same or is no existence of any opinion words in the text. Result of sentiment analysis are shown in Table 2 showing the distribution of positive, negative and neutral tweets of every Government. Which indicate the how polarity varies with different government schemes on twitter dataset.

Table 2. Result of Polarity of Twitter Dataset

Trend Name	Positive Sentiments	Negative Sentiments	Neutral Sentiments	Tweets used
#GST4India	2427	123	2450	5000
#startupindia	924	192	1384	2500
#SwachBharat	1983	409	2138	4530
#FinBudget	7791	4662	12547	25000
#Digital india	1238	411	3351	5000
#RailwayBudget	1257	618	2115	4000
#MakeInIndia	3383	454	6163	10000
#Kashmir	923	1399	1678	4000
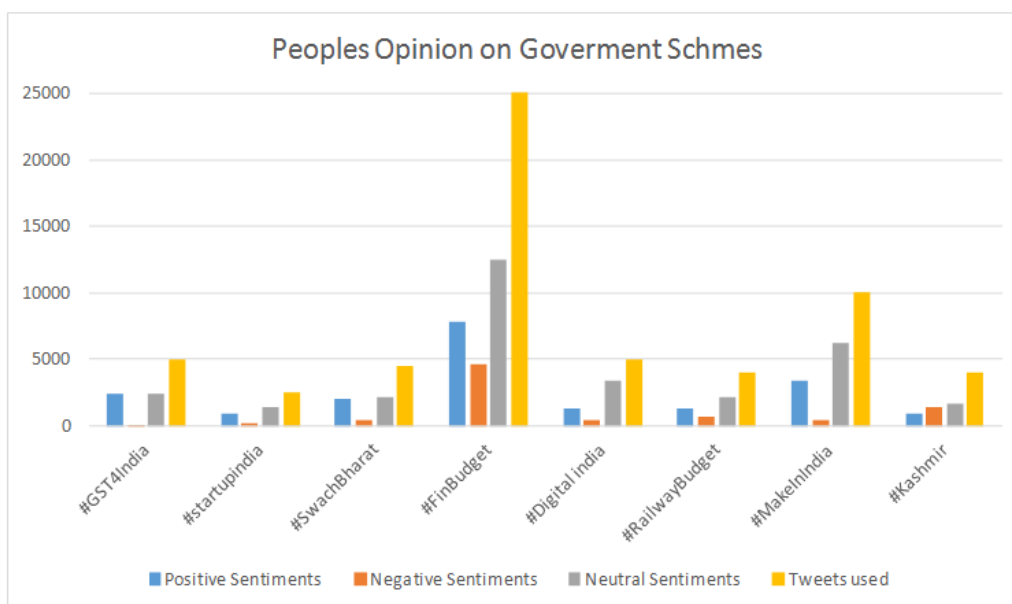
Graph 1 No. of Positive, Negative and Negative Tweets vs. #Hashtag

VI. Conclusions And Future Work

In this experimental work, we used Twitter API using R tool which is open source. Tweets from twitter has been collected and gives to pre-processing task in that tool. R open source tool is used in text mining and also to crawl streaming data from social media like twitter and Facebook etc. Government schemes data are pre-processed in R tool for sentiment analysis and opinion mining. There are different supervised and unsupervised approaches and different lexicons, dictionaries and corpus based methods which are very helpful in Sentiment Analysis. Different dataset are available for Government Schemes. In this paper sentiment score has been calculated and counted number of positive, negative and neutral tweets for given #Hashtag and can predict the public opinion of particular Scheme. As per above analysis of different #Hashtags tweets for sentiment analysis, peoples and governments can find the public opinion behind that declared scheme. Table of summary shows the used methods and dataset for particular Schemes. Future work about government schemes and sentiment analysis is find out aspects and their polarity of the schemes which helps for implantation of government schemes effectively to take decision to upcoming scheme regarding public satisfaction.

References

[1]. Rajkumar S. Jagdale, Vishal S. Shirsat, Sachin N. Deshmukh," Sentiment Analysis of Events from Twitter Using Open Source Tool International Journal of Computer Science and Mobile Computing ISSN 2320–088X IMPACT FACTOR: 5.258 IJCSMC, Vol. 5, Issue. 4, April 2016, pg.475 – 485

[2]. Xing Fang and Justin Zhan."Sentiment analysis using product review data". Journal of Big Data 2015. DOI: 10.1186/s40537-015-0015-2

[3]. Seyed-Ali Bahrainian, Andreas Dengel, "Sentiment Analysis and Summarization of Twitter Data", CSE, 2013, 2013 IEEE 16th International Conference on Computational Science and Engineering (CSE), 2013 IEEE 16th International Conference on Computational Science and Engineering (CSE) 2013, pp. 227-234, doi:10.1109/CSE.2013.44

[4]. Prashant Raina, "Sentiment Analysis in News Articles Using Sentic Computing", ICDMW, 2013, 2013 IEEE 13th International Conference on Data Mining Workshops (ICDMW), 2013 IEEE 13th International Conference on Data Mining Workshops (ICDMW) 2013, pp. 959-962, doi:10.1109/ICDMW.2013.27.

[5]. Geetika Gautam and Divakar yadav, "Sentiment Analysis of Twitter Data Using Machine Learning Approaches and Semantic Analysis", 978-1-4799-5172-7/, pp.437-442 2014 IEEE.

[6]. Federico Neri , Carlo Aliprandi , Federico Capeci , Montserrat Cuadros , Tomas By, Sentiment Analysis on Social Media, Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012), p.919-926, August 26-29, 2012

[7]. Godbole, N.; Srinivasaiah, M. & Skiena, S. (2007), Large-Scale Sentiment Analysis for News and Blogs, in 'Proceedings of the International Conference on Weblogs and Social media (ICWSM)'.

[8]. H. Tang, S. Tan, and X. Cheng. A Survey on Sentiment Detection of Reviews. Expert Systems with Applications. Elsevier. 2009, doi:10.1016/j.eswa.2009.02.063.

[9]. Yulan He,"Learning sentiment classification model from labeled features. "Proceedings of the 19th {ACM} Conference on Information and Knowledge Management, {CIKM} 2010, Toronto, Ontario, Canada, October 26-30, 2010.

[10]. Long-Sheng Chen, Cheng-Hsiang Liu, Hui-Ju Chiu, "A neural network based approach for sentiment classification in the blogosphere", Journal of Informetrics 5 (2011) 313–322.

[11]. ZHU Jian , XU Chen, WANG Han-shi, " Sentiment classification using the theory of ANNs", The Journal of China Universities of Posts and Telecommunications, July 2010, 17(Suppl.): 58–62 .

[12]. Bo Pang, Lillian Lee.A Sentimental Education: Sentiment Analysis Using Subjectivity. Proceedings of ACL, pp. 271--278, 2004

[13]. Bai, R. Padman, "Markov blankets and meta-heuristic search: Sentiment extraction from unstructured text," Lecture Notes in Computer Science, vol. 3932, pp. 167–187, 2006.

[14]. Kennedy and D. Inkpen, "Sentiment classification of movie reviews using contextual valence shifters," Computational Intelligence, vol. 22, pp. 110–125, 2006.

[15]. Zhou, L. and P. Chaovalit (2008), Ontology-supported Polarity Mining, Journal of the American Society for Information Science and Technology, 59(1), 1-13.

[16]. Rudy Prabowo, Mike Thelwall, "Sentiment analysis: A combined approach." Journal of Informetrics 3 (2009) 143–157.

[17]. Rui Xia, Chengqing Zong, Shoushan Li, "Ensemble of feature sets and classification algorithms for sentiment classification", Information Sciences 181 (2011) 1138–1152.

[18]. Popescu, A. M., Etzioni, O.: Extracting Product Features and Opinions from Reviews, In Proc. Conf. Human Language Technology and Empirical Methods in Natural Language Processing, Vancouver, British Columbia, 2005, 339–346.

[19]. Qingliang Miao, Qiudan Li, Ruwei Dai, "AMAZING: A sentiment mining and retrieval system", Expert Systems with Applications 36 (2009) 7192–7198.

[20]. Turney, Peter (2002). "Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews". Proceedings of the Association for Computational Linguistics. pp. 417–424.

[21]. Hatzivassiloglou, V., & McKeown, K.R. 1997. Predicting the semantic orientation of adjectives. Proceedings of the 35th Annual Meeting of the ACL and the 8th Conference of the European Chapter of the ACL (pp. 174-181). New Brunswick, NJ: ACL.

[22]. Tong, R.M. 2001. An operational system for detecting and tracking opinions in on-line discussions. Working Notes of the ACM SIGIR 2001 Workshop on Operational Text Classification (pp. 1-6). New York, NY: ACM.

[23]. Wiebe, J.M. 2000. Learning subjective adjectives from corpora. Proceedings of the 17th National Conference on Artificial Intelligence. Menlo Park, CA: AAAI Press.

[24]. Pang, Bo; Lee, Lillian (2005). "Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales". Proceedings of the Association for Computational Linguistics (ACL). pp.115–124.

[25]. Yang Liu, Xiangji Huang, Aijun An, Xiaohui Yu (2007). "ARSA: A Sentiment-Aware Model for Predicting Sales Performance Using Blogs" SIGIR'07, July 23–27, 2007, Amsterdam, The Netherlands.

[26]. R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[27]. Cataldo Musto, Giovanni Semeraro, Marco Polignano."A comparison of Lexicon-based approaches for Sentiment Analysis of microblog posts".8th International Workshop on Information Filtering and Retrieval Pisa (Italy) December 10, 2014.

[28]. Andrea Esuli Baccianella, Stefano and Fabrizio Sebastiani. SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In Proceedings of LREC, volume 10, pages 2200-2204, 2010.

[29]. Erik Cambria and Amir Hussain. Sentic computing. Springer, 2012.

[30]. Kamps, Maarten Marx, Robert J. Mokken and Maarten De Rijke, "Using wordnet to measure semantic orientation of adjectives", Proceedings of 4th International Conference on Language Resources and Evaluation, pp. 1115-1118, Lisbon, Portugal, 2004.

[31]. Stone, P. J., Dunphy, D. C., Smith, M. S., and Ogilvie, D. M. (1966). The General Inquirer: A Computer Approach to Content Analysis. MIT Press.

[32]. Clayton J. Hutto, Eric Gilbert."VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text.".Proceedings of the Eighth International Conference on Weblogs and Social Media,ICWSM 2014, Ann Arbor, Michigan, USA, June 1-4, 2014.

[33]. Kaiquan Xu , Stephen Shaoyi Liao , Jiexun Li, Yuxia Song, "Mining comparative opinions from customer reviews for Competitive Intelligence", Decision Support Systems 50 (2011) 743–754.

[34]. Long-Sheng Chen, Cheng-Hsiang Liu, Hui-Ju Chiu, "A neural network based approach for sentiment classification in the blogosphere", Journal of Informetrics 5 (2011) 313–322.

[35]. Hanhoon Kang, Seong Joon Yoo , Dongil Han," Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews". Expert Systems with Applications 39 (2012) 6000–6010.

[36]. Lin Y, Zhang J, Wang X, Zhou A (2012) An information theoretic approach to sentiment polarity classification. In: Proceedings of the 2Nd Joint WICOW/AIRWeb Workshop on Web Quality,WebQuality '12. ACM, New York, NY, USA.pp 35–40

[37]. Hu, Minqing and Bing Liu. Mining and summarizing customer reviews. In Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004). 2004.

[38]. Wiebe, Janyce, Rebecca F. Bruce, and Thomas P. O'Hara. Development and use of a gold-standard data set for subjectivity classifications. In Proceedings of the Association for Computational Linguistics (ACL-1999). 1999.

[39]. Emitza Guzman, Walid Maalej, "How Do Users Like This Feature? A Fine Grained Sentiment Analysis of App Reviews", RE, 2014, 2014 IEEE 22nd International Requirements Engineering Conference (RE), 2014 IEEE 22nd International Requirements Engineering Conference (RE) 2014, pp. 153-162, doi:10.1109/RE.2014.6912257

[40]. Chau, M., & Xu, J. (2007). Mining communities and their relationships in blogs: A study of online hate groups. International Journal of Human – Computer Studies, 65(1), 57–70.

[41]. Lars Kai Hansen, Adam Arvidsson, Finn Årup Nielsen, Elanor Colleoni,Michael Etter, "Good Friends, Bad News - Affect and Virality inTwitter", The 2011 International Workshop on Social Computing, Network, and Services (SocialComNet 2011).

[42]. Pang, Bo; Lee, Lillian; Vaithyanathan, Shivakumar (2002). "Thumbs up? Sentiment Classification using Machine Learning Techniques". Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP). pp. 79–86.