Twitter Sentiment Analysis: A Review

Kishori K. Pawar, Pukhraj P Shrishrimal, R. R. Deshmukh

Abstract— The basic knowledge required to do sentiment analysis of Twitter is discussed in this review paper. Sentiment Analysis can be viewed as field of text mining, natural language processing. Thus we can study sentiment analysis in various aspects. This paper presents levels of sentiment analysis, approaches to do sentiment analysis, methodologies for doing it, and features to be extracted from text and the applications. Twitter is a microblogging service to which if sentiment analysis done one has to follow explicit path. Thus this paper puts overview about tweets extraction, their preprocessing and their sentiment analysis.

Index Terms— CIssifier, Lexicons, Machine Learning, Opinion Mining, Sentiment Analysis ,Twitter, Text Mining.

1 Introduction

IN text mining, Sentiment Analysis and Opinion Mining consists study of sentiments, attitudes, reactions, evaluation of the content of the text. Many times while analyzing peoples' opinions, sentiments, evaluations, attitudes, reactions towards entities, such as services, products, organizations, individuals , events, topics, issues and their attributes Sentiment Analysis is also called as Opinion Mining [1].

Twitter is a microblogging media in real time to express the persuasion of a person or group about a particular topic to appear going on a timeline. The message which is displayed on Twitter is named as Tweet. The users are made by friends and followings, tweets and their timeline are key components of Twitter. The chronologically sorted collection of multiple tweets is the timeline. A person can express his view in front of the world in various *forms* like multimedia, text etc. Because of popularity of Twitter as an information source, it led to development of applications and research in many spheres. Twitter is used in predicting the happenings of earthquakes and identifying relevant users to follow to obtain disaster relevant information [2].

Web search applications, Real world applications like current trends in world, world events, extracting latest information about incidents uses micro blog data for their analysis and conclusion making [3].

Twitter is much different from other social media. Microblogging is one vital fact and it is more opinion oriented than informative about a topic. There are many abbreviations, symbols, and the content is many times similar to conversation. Mutual Acceptance from both users is not needed in Twitter because of its Asymmetric Following Model [4].

Thus huge and varied amount of knowledge can be extracted from the tweets. This paper studies about the sentiment detection from the tweets. The step by step methodology and the comparative analysis of existing system to do sentiment classification is discussed in this review paper.

2 Sentiment Analysis

2.1 Definition

Sentiment and Subjectivity are mainly context and domain dependent. Not only the changes in vocabulary are the reason behind that but one more reason is the dual meaning or sentiments of same expression in different domains. Consider the example of expression 'go and read the book'. In case of book reviews this expression gives the positive polarity about the product but in case of movie review the same expression gives negative polarity about the product. Sentiment Analysis is more focused on extraction of polarity about a particular topic rather than assigning a particular emotion to the text.

Opinion Mining and Sentiment Analysis are the branches of Text Mining which refer to the process of extracting nontrivial patterns and interesting information from unstructured script documents. We can say that they are the addition to data mining and knowledge discovery. Opinion Mining and Sentiment Analysis focus on polarity detection and emotion recognition correspondingly. Opinion Mining has more marketable potential higher than data mining as it the most natural form of storing the information in text format. It is much complex task than data mining because it has to deal with unstructured and fuzzy data. It is a multi-disciplinary area of research because it constitutes adoption of techniques in information retrieval, text analysis and extraction, auto-categorization, machine learning, clustering, and visualization.

Though Sentiment Analysis and Opinion Mining might look the same as the fields like traditional text mining or fact based analysis, it varies because of following facts. Sentiment Classification is the binary polarity classification which deals with a relatively small number of classes. Sentiment classification is easy task compared to text auto-categorization. While Opinion mining exhibits many additional tasks other than sentiment polarity detection like summarization and all.

2.2 Levels

We can divide sentiment analysis in following levels. [5]

2.2.1 Document

The task at this level is classifying the sentiment for document. The document is on single topic is considered. Thus texts which comprise comparative learning cannot be considered under this level.

2.2.2 Sentence

The task at this level goes to the sentences; it determines whether each sentence expresses a positive opinion, negative opinion, or neutral view. If a sentence states no opinion means it is a neutral. This level of analysis is closely related to subjectivity classification. The subjective statement displays the polarity of an entity in affirmative-negative terms i.e. good-bad terms. Hence it is easy to obtain sentiment from it. But Objective statement does not give separation directly by affirmative-negative terms. These are abstract sentences which are fact based.

2.2.3 Entity or Aspect

Aspect level gives detailed analysis. The core task of entity level is to identification of aspect of the text [1]. For example in a review of mobile if a customer says," Sound is good but the handset is not handy." In this review the aspect are sound and handiness. Here sentiment analysis becomes two level task i.e. finding the aspects in the text and then classifying in respective aspect. Aspect level sentiment analysis is superior to Document and Sentence level sentiment analysis. Sentiment analysis of topic or body which may or may not be hidden in the document is done. Thus comparative statements are also part of entity level sentiment analysis [6].

2.3 APPROACHES

We can do opinion mining and sentiment analysis in fol-lowing ways: keyword spotting, lexical affinity, statistical methods.

2.3.1 Keyword Spotting

In this technique the text is categorized based on the pres-ence of fairly unambiguous words present in it. Thus the words or keywords present in the text have the importance with respect to sentiment analysis.

2.3.2 Lexical affinity

For a particular emotion, Lexical affinity assigns arbitrary words a probabilistic similarity.

2.3.3 Statistical methods

It calculates the valence or target of affective keywords and word co-occurrence frequencies on the base of a large training corpus.

In early work it was aim to classify entire document into overall affirmative or negative. These systems mainly depend on supervised learning approaches which depend on manually labeled data. The examples of such systems are movie or product review databases. Many times sentiments are not restricted to document level texts. It can be extracted from sentence level text. In such cases sentiment analysis can be done using detected opinion-bearing lexicon items. Or sentiments are not limited to particular target, they can contrary towards same topic or multiple topics can be present in the same document [7]..

2.4 Features

Sentiment features are as follows:

2.4.1 Terms presence and frequency

These features are nothing but individual words or word n-grams and their frequency counts. It either uses the term frequency weights or gives binary weighting to the words.

2.4.2 Parts of speech (POS)

It set up finding adjectives from the text, as they are important indicators of opinions.

2.4.3 Opinion words and phrases

These words themselves express opinion about the product or service in the text. For e.g. good or bad, like or hate. Some phrases also express opinions without using opinion words.

Negations: the presence of negative words may change the opinion orientation like not good is equivalent to bad.

2.5 Methodologies

Sentiment Classification techniques can be roughly divided into Lexicon based approach, Machine Learning approach and hybrid approach. The Machine Learning Approach (ML) applies the famous ML algorithms and it uses linguistic features. The Lexicon-based Approach depends on a sentiment lexicon. Lexicon is a collection of known and precompiled sentiment terms. It is again divided into dictionary-based approach and corpus-based approach which use semantic or statistical methods to find sentiment polarity of the text. The Hybrid Approach combines both approaches and it is very common with sentiment lexicons playing a key role in the majority of methods.

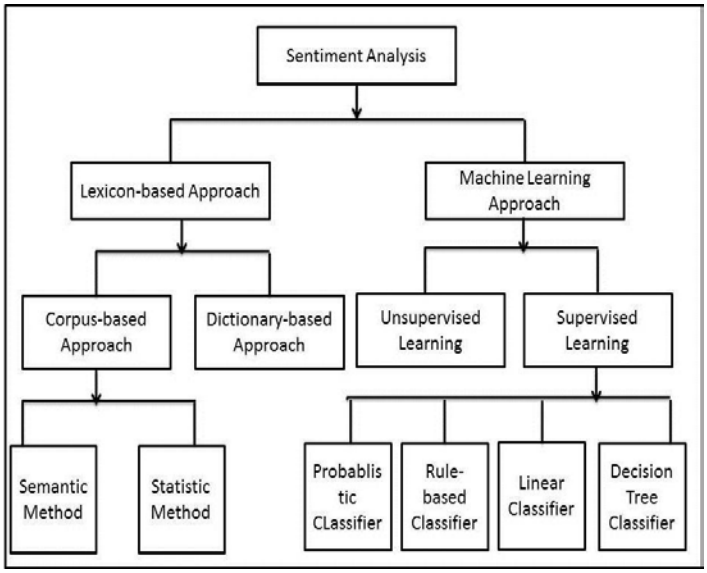

Fig 1 Sentiment Analysis Techniques [8]

2.6 Methodologies

Sentiment analysis applications have range to almost every possible domain, from consumer products, ser-

vices, and commercial services to societal events and political elections. Nearly all companies need Sentiment Analysis and Opinion Mining for different applications in different scenarios. In many product review websites like Yelp, Epinions reviews and feedbacks are explicitly asked in their web interfaces. Sentiment Analysis is not only limited to product reviews but expands it wing to many fields like political/governmental issues. Opinion Mining can increase capabilities of Customer Relationship Management (CRM) and Recommendation Systems by collecting positive and negative sentiments of the consumer. By using Sentiment Analysis techniques wired systems displaying advertisements can detect web pages that contain sensitive content inappropriate for trailers placement. Companies are applying different marketing strategies like collecting opinions of general public about the products and services. These sentiments can be mined using Sentiment Analysis for Business Intelligence. Not only the commercial market but government intelligence also uses opinion mining to monitor the negative communications over social media.

3 TWITTER

3.1 Definition

The word 'micro' in microblogging specifies the limitation of content of the opinion expressed on it. A twitter user can compose at max 140 characters per each tweet. A tweet is not only a simple text message but it is a combination of text data and Meta data associated with the tweet. These attributes are the features of tweets. They expresses the content of the tweet or what is that tweet about. The Metadata can be utilized to find out the domain of the tweet. The Metadata of tweet are some entities and places. These entities include user mentions, hashtags, URLs, and mediaUsers, Twitter userID. RT stands for retweet, '@' followed by a user identifier report the user, and '#' followed by a word characterizes a hashtag. Work on the Twitter in this paper is limited up to text data.

3.2 Twitter Features

For Opinion Retrieval following features can be useful:

1) Twitter Specific Features

a) URL

Many tweets share a link along with the introduction to the links. The sharing link is initiated as URL. Presence of URL, gives its feature value as 1, else is 0.

b) Mention

In a tweet when user want to refer to another user he can write his name starting with @ symbol. It is called as Mention and it also represented as "@username". If tweet encloses mention the binary feature representing it will have value 1, else is 0.

c) Recency

When the query is fired to get a tweet, it is better to get most recent tweet about that matter. Thus Recency feature measures the age of tweet in seconds after its generation.

d) Hashtag

It is a word starting with # symbol. It refers to a word about the content of text or indicating the topic of tweet. The binary feature value gives the answer of whether the tweet contains hashtag or not.

e) Emoticons

These are facial expressions pictorially characterized using punctuation and letters; they express the user's mood.

f) Retweet

A tweet can be just a statement made by a user, or could be a reply to another tweet. Retweets are marked with either "RT" followed by '@user id' or "via @user id". Retweet is considered the feature that has made Twitter a new medium of information dissemination as well as direct communication.

g) Singleton

If a tweet has no reply or a retweet, then the tweet is called as singleton.

2) Author features

As twitter is a social network,the rich author information can also be used for the opinion retrieval task.

a) Followers and Friends

In Twitter the relation between two users is as follows : Consider two users A and B. User B can follow user A. In such case user B is a follower of user A and user A is a friend of user B. The number of followers shows the popularity of a person. The tweet or status posted by user A will be streamed to user B's profile as he is follower of A.

b) Statuses

The number of tweets posted by an author shows the activeness of the author. For measuring the Tweets Ranking this feature can be used.

c) Listed

According to a topic or a social relationship, a user can group his friend into different lists. If the tweets are interesting to large number of population then the user is listed many times. Thus this feature can be used for tweet ranking [9].

4 SENTIMENT ANALYSIS IN TWITTER

Sentiment analysis is all about extracting opinion from the text. There are various aspects, reasons, orientation of extracting these emotions according to reason behind the analysis. Event detection, location detection etc. tasks can be done on tweets. When this task is accomplished on twitter data, the framework or architecture to do sentiment analysis varies according to what type of result one want to achieve from the tweets. One more important factor behind the varying nature of flow of twitter sentiment analysis is use of different methodolo-

gies and techniques. Many times researchers derive their own framework or flow to do sentiment analysis to increase efficiency of the result. Some of common steps in twitter sentiment analysis and the keywords in it are defined below:

4.1 Preprocessing

Despite of these generalized orientation of framework of twitter sentiment analysis, we can frame up this topic into the following workflow. Thus the generalized steps involved in this framework are as follows:
Before starting sentiment analysis, the data preprocessing need to be done.

4.1.1 Removal of Non-English Tweets
When the tweets are extracted from big datasets like TREC or Clueweb dataset, it contains English as well as non-English tweets. Therefore, we have to run language identification on each tweet, and have to delete from our collection all tweets that are assigned a 0-probability of being English.

4.2 Feature Selection

4.2.1 Lexicon Features
Based on the subjectivity of the word we can classify the words into positive, negative and neutral lexicons. We have to compare each word with predefined wordnet libraries.

4.2.2 Part-of-speech Features
Parts-of speech features i.e. nouns, adverbs, adjectives, etc. in each tweets are tagged.
4.2.3 Micro-blogging Features
By creating binary features we can detect the presence of positive, negative, and neutral emotions. By the presence of abbreviations and intensifiers we can classify tweets in positive, negative and neutral. Online available slang dictionaries can be used for emotions and abbreviations [11].
4.2.4 Steps to Extract Features

4.2.4.1 Case Normalization
In this step entire document is converted into lowercase.
4.2.4.2 Tokenization
Tokenization is splitting up the systems of text into personal terms or tokens. This procedure can take many types, according to the terminology being examined. For English, effective tokenization technique is to use white space and punctuation as token delimiters.
4.2.4.3 Stemming (Snowball)
Stemming is the procedure of decreasing relevant tokens into a single type of token. This procedure contains the recognition and elimination of suffixes, prefixes, and unsuitable pluralizations.
4.2.4.4 Generate n-Grams
Character n-grams are 'n' nearby figures from a given feedback sequence. For example, a 3- gram of a phrase 'FORM' would be '_ _ F', '_FO', 'FOR', 'ORM', 'RM_', 'M_ _'. N-grams of dimension one are known as 'unigram', two dimensional grams are known as 'bigram', three-dimensional grams are known as 'trigram'. And for the rest dimensions it is called as n-grams [11].

4.1.2 Removal of Re-tweets
We have to delete any text that followed an RT token (as well as the RT token itself), since such text typically corresponds to quoted (retweeted) material.

4.1.3 Conversion to ASCII
Many tweets contain unusual or non-standard characters, which can be problematic for down-stream processing. To address these issues, we have to use a combination of BeautifulSoup5 and Unidecode6 to convert and transliterate all tweets to ASCII.

4.1.4 Removal of Empty Tweets
After completing all of the other pre-processing, we have to delete any empty tweets.

4.1.5 Restoration of Abbreviations
We can restore popular abbreviations used in the tweets, to their corresponding original forms using a lexicon of abbreviations (e.g. "wknd" to "weekend").Punctuations are kept since people often express sentiment with tokens such as ":)", ":-)". These emotions can also be used for sentiment classification [10].

4.3 Comparative Analysis of Different Workflows

The table below shows works to be accomplished by various authors on Twitter sentiment analysis. What are the various sources of tweet dataset, what is the methodology? Alongwith the advancements as well as drawback of the techniquesa are discussed below:

TABLE 1. COMPARATIVE ANALYSIS OF TECHNIQUES, APPROACHES OF SENTIMENT ANALYSIS

Sr. No.	Author & Year of Publication	Discussion	Accuracy/ Sentiment Score/Analysis
1	Dhiraj Gurkhe et al.(2014)[12] Dataset	Twitter Training Data: Amalgamation of different labeled dataset(Movie Review Datasets, Sanders-Twitter Sentiment Corpus, Twitter data based on emoticons, Sentiment Lexicon)	Accuracy: Unigram :81.25% Bigram :15% Uni+Bigram

			Test Data: Twitter Dataset using API	:67.50%
		Technique	Approach: Machine Learning Unigram, Bigram, Uni+Bigram features with Naïve Bayes	
		Advancements	It gives best results with Unigram detection without neutral labels.	
		Drawback	Size of training data is less leading to have less accuracy. Sarcasm cannot be detected;leading to false classification.	
2	Riya Suchdev et al.(2014)[13]	Dataset	Sanders Analytics dataset	Accuracy: 100%
		Technique	Machine Learning & knowledge-based approach using feature vector	
		Advancements	Use of hybrid approach gives 100% of accuracy.	
3	Eman M.G. Younis et al.(2015)[14]	Dataset	Twitter Dataset using Twitter API(2000 tweets for Tesco and 803 tweets for Asda)	Sentiment Score: Tesco- 0.1595 Asda -0.00373599
		Technique	Lexicon-based approach	
		Advancements	Sentiment Analysis can be carried out by unsupervised technique without using preclassified training dataset. System gives a way to improve business competitive and customer relationship management in real-time.	
		Drawback	Sarcasm cannot be detected; leading to false classification.	
4	Subhabrata Mukherjee et al. [15]	Dataset	Dataset1: Tweets crawled using Tweet Fetcher module(8507 tweets on 20 different domains) Dataset 2: Using Twitter API set of 15,214 tweets based on hashtags	Accuracy Dataset1: 2 class: 66.69% 3 class: 56.17% Dataset2: 2 class:v88.53%
		Technique	Hybrid Approach: machine learning approach & Rule based approach with extended module for each phase in architecture .	
		Advancements	TwiSent achieves a higher negative precision improvement than positive precision improvement over C-Feel-It, which indicates it can capture negative sentiment strongly.	
		Drawback	System cannot capture sarcasm or implicit sentiment. Supervised system accuracy suffers due to sparse feature space due to inherent text limit of tweets	
5	Alexander Pak et al. [16]	Dataset	Collected using Twitter API	Accuracy vs. decision graph is plotted
		Technique	Multinomial Naive Bayes classifier on n-grams and part-of-speech & statistical linguistic analysis of corpus	
		Advancements	The best performance is achieved when using bigrams	
6	Efthymios Kouloumpis et al.(2011)[11]	Dataset	1)Edinburgh Twitter corpus 2)emoticon data set (EMOT) 3) iSieve data set	Accuracy (For model which uses n-grams+lex+twit) Dataset1:0.74 Dataset2:0.75 Dataset1+Dataset2:0.75
		Technique	Approach: Lexicon based Features: 1) n-grams 2) POS 3) lexs 4) twits	
		Drawback	They experimented the system with SVMs, which gave similar trends, but lower results overall.	

7	Akshat Bakliwal et al. (2013)[17]	Dataset	Twitter search API between 20th and the 25th of January 2011(7,916 tweets)	Accuracy: Lexicon based approach:58.9% Lexicon based and machine learningapproach: 61.6%.
		Technique	Approach: Hybrid (Unsupervised lexicon-based approach and supervised machine learning) Method: SVMLight	
		Advancements	The accuracy increases slightly when the lexicon-based information is encoded as features and employed together with bag-of-word features in a supervised machine learning setup.	
		Future scope	More experiments need to be done to classify sarcastic tweets	
8	Akshi Kumar et al.(2012)[18]	Dataset	Collected using Twitter API	Tweet is classified as Positive (SS>0) Negative (SS<0) Neutral (SS=0)
		Technique	Lexicon based approach (corpus-based method &dictionary-based method). Sentiment score(SS) is calculated.	
9	I.Hemalatha et al. (2013)[19]	Dataset	Manrnual collection of tweets from website of social network sites	Accuracy: Not mentioned Classification: 52(positive) 32(nagative) 16(neutral)
		Technique	Approach: Machine learning supervised learning Algorithm: Naive Bayes and Maximum Entropy algorithm	
10	Namita Mittal et al. (2013)[20]	Dataset	Labelled dataset of Stanford University consisting of 60,000 tweets	Accuracy: SWN:66.052 PBM: 71.625 SWN then PBM: 70.193 PBM then SWN: 72.35 Hybrid:73.72
		Technique	Approach: Hybrid Methods: SentiWordNet (SWN) and probability based method(PBM) to determine the polarity of tweets	
		Advancements	Effectiveness of the proposed hybrid approach for sentiment analysis is greater than individual	
		Future scope	Handling sarcasm and more discourse relations.	
11	Christopher Johnson et al.[21]	Dataset	1) Emoticon dataset consists of 1.6 billion tweets 2) Using the Apache Hadoop framework, 550k tweets are filtered, containing the case-insensitive string `obama' from a 50GB corpus consisting of tweets gathered between September 2009 and October 2010. 3) They used the Gallup poll's data on President Obama's job approval from the same time period as our Obama dataset (2009 to 2010).	Accuracy: LR:43 ME:60 LP with no Retweet :75 LP with Retweet: 67 LP with no Retweet/Reply edges and seeded hashtags: 78
		Technique	Approach: Rrule-based, supervised, and semi-supervised learning methods. Methods: Lexical Ratio, Maximal Entropy Classifier, Label Propogation	
		Advancements	Twitter specific features such as hashtags, specific to the topic one is analyzing, can be used to improve the accuracy of sentiment predictions is shown in this paper.	
12	Lei Zhang et al.(2011)[22]	Dataset	Collected using Twitter API	Accuracy 0.854
		Technique	Approach: Machine learning and Lexicon based Method : Support Vector Machines	
13	Dmitry Davidov	Dataset	1) Twitter dataset generously provided to us by Brendan O'Connor. This dataset includes over 475 million tweets.	Average harmonic f-score for senti-

	et al. (2010)[23]		2) Amazon Mechanical Turk (AMT) service inorder to obtain a list of the most commonly used and unambiguous ASCII smileys	ment hashtags : 0.8
		Technique	Supervised sentiment classification using k-nearest neighbors Evaluation using cross-validation	Average harmonic f-score for smileys: 0.86
		Advancements	They proposed two different methods which allow an automatic identification of sentiment type overlap and inter-dependencies.	
		Future scope	Automated clustering of sentiment types and sentiment dependency rules need to be done	
14	Sunil B. Mane et al. (2014)[24]	Dataset	Collection of http://www.cs.tau.ac.il/~kfirbar/mlproject/twitter.data	Accuracy: 72.27
		Technique	Naïve-Bayes classifier	
		Advancements	Lower response time has achieved by use of Hadoop framework.	

5 Conclusion

Thus the basic knowledge required to do sentiment analysis of Twitter is well stated in this review paper. What is Sentiment Analysis with respect to levels of sentiment analysis, what are the approaches to do sentiment analysis, methodologies for sentiment analysis, features to be extracted from text and the applications where it can be utilized is mentioned hierarchically. If we want to do Twitter's sentiment analysis we need to know about the twitter, about extracting the tweets, its structure, their meaning. This paper gives brief notion of tweets. When one wants to do sentiment analysis of tweets, he has to do it in a specialized aspect of sentiment analysis. So the brief knowledge about Twitter Sentiment Analysis is given in this paper. Different methods and techniques are discussed in a comparative manner. The accuracy/ result of each method enables us to imagine the efficiency of applied technique in respective circumstances.

Acknoeledgement

This review has been partially supported by Department of Computer Science and IT Dr.Babasaheb Ambedkar Marathwada University, Aurangabad. The views expressed here are those of the authors only.

References

[1] Bing Liu, Sentiment analysis and opinion mining, Synthesis Lectures on Human Language Technologies 5 (2012), no. 1, 1-167.

[2] Shamanth Kumar, Huan Liu, Fred Morstatter, , "Twitter Data Analytics", Morgan & Claypool Publishers, Springer, 2014.

[3] Peter Korenek, Marián Simko, "Sentiment Analysis on Microblog Utilizing Appraisal Theory", Published at Springer pp. 667–702, 2013.

[4] Matthew A. Russell, "Mining the Social Web, Second Edition", 2nd Edition of O'Reilly Media, May 2013.

[5] Dudhat Ankitkumar, Prof. R. R. Badre, Prof. Mayura Kinikar, "A Survey on Sentiment Analysis and Opinion Mining", International Journal of Innovative Research in Computer and Communication Engineering Vol. 2, Issue 11, November 2014.

[6] Walaa Medhat, Ahmed Hassan, Hoda Korashy, "Sentiment analysis algorithms and applications: A survey", Ain Shams Engineering Journal , 1093–1113, 2014.

[7] Erik Cambria, Amir Hussain, "Sentic Computing Techniques, Tools, and Applications", Published at Springer May 9, 2012.

[8] Rupali P. Jondhale,Manisha P. Mali"Study on Distinct Approaches for Sentiment Analysis", International Journal of Computer Applications (0975 – 8887) Volume 111 – No 17, February 2015.

[9] Zhunchen Luo, Miles Osborne, Ting Wang, "An effective approach to tweets opinion retrieval", Published at Springer, World Wide Web DOI, 2013.

[10] Lei Zhang, Riddhiman Ghosh, Mohamed Dekhil, Meichun Hsu, Bing Liu, "Combining Lexicon-based and Learning-based Methods for Twitter Sentiment Analysis", HP Laboratories HPL-2011-89.

[11] Efthymios Kouloumpis, TheresaWilson, Johanna Moore, "Twitter Sentiment Analysis: The Good the Bad and the OMG!", Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media.

[12] Dhiraj Gurkhe, Niraj Pal, Rishit Bhatia, "Effective Sentiment Analysis of Social Media Datasets using Naive Bayesian Classification", International Journal of Computer Applications (0975 8887), Volume 99 - No. 13, August 2014.

[13] Riya Suchdev, Pallavi Kotkar,Rahul Ravindran, "Twitter Sentiment Analysis using Machine Learning and Knowledge-based Approach", International Journal of Computer Applications (0975 – 8887),Volume 103 – No.4, October 2014

[14] Eman M.G. Younis,"Sentiment Analysis and Text Mining for Social Media Microblogs using Open Source Tools: An Empirical Study", International Journal of Computer Applications (0975 – 8887), Volume 112 – No. 5, February 2015.

[15] Subhabrata Mukherjee, Akshat Malu, Balamurali A.R., Pushpak Bhattacharyya, "TwiSent: A Multistage System for Analyzing Sentiment in Twitter",

[16] Alexander Pak, Patrick Paroubek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining",

[17] Akshat Bakliwal, Jennifer Foster, Jennifer van der Puil, Ron O'Brien, Lamia Tounsi and Mark Hughes, "Sentiment Analysis of Political Tweets: Towards an Accurate Classifier ", Proceedings of the Workshop on Language in Social Media (LASM 2013), pages 49–58,Atlanta, Georgia, June 13 2013.

[18] Akshi Kumar, Teeja Mary Sebastian, "Sentiment Analysis on Twitter", IJCSI International Journal of Computer Science Issues, (1694-0814), Vol. 9, Issue 4, No 3, July 2012.

[19] I.Hemalatha, Dr. G. P Saradhi Varma, Dr.Govardhan,"Sentiment Analysis Tool using Machine Learning Algorithms", International Journal of Emerging Trends & Technology in Computer Science (IJETTCS),Volume 2, Issue 2, March – April 2013.

[20] Namita Mittal, Basant Agarwal, "A Hybrid Approach for Twitter Sentiment Analysis", Proceedings of ICON-2013: 10th International Conference on Natural Language Processing, pp: 116-120, 2013.

[21] Christopher Johnson, Parul Shukla, Shilpa Shukla, "On Classifying the Political Sentiment of Tweets",

[22] Lei Zhang, Riddhiman Ghosh, Mohamed Dekhil, Meichun Hsu, Bing Liu, "Combining Lexicon-based and Learning-based Methods for Twitter Sentiment Analysis", Copyright 2011 Hewlett-Packard Development Company, L.P.

[23] Dmitry Davidov, Oren Tsur,Ari Rappoport, "Enhanced Sentiment Learning Using Twitter Hashtags and Smileys",Coling 2010: Poster Volume, pages 241–249, Beijing, August 2010.

[24] Sunil B. Mane, Yashwant Sawant, Saif Kazi, Vaibhav Shinde, " Real Time Sentiment Analysis of Twitter Data Using Hadoop", International Journal of Computer Science and Information Technologies, (3098 - 3100),Vol. 5 (3) , 2014.