Sentiment Analysis on Twitter

Akshi Kumar and Teeja Mary Sebastian

**Department of Computer Engineering, Delhi Technological University
Delhi, India**

Abstract

With the rise of social networking epoch, there has been a surge of user generated content. Microblogging sites have millions of people sharing their thoughts daily because of its characteristic short and simple manner of expression. We propose and investigate a paradigm to mine the sentiment from a popular real-time microblogging service, Twitter, where users post real time reactions to and opinions about "everything". In this paper, we expound a hybrid approach using both corpus based and dictionary based methods to determine the semantic orientation of the opinion words in tweets. A case study is presented to illustrate the use and effectiveness of the proposed system.

Keywords: Microblogging, Twitter, Sentiment Analysis

1. Introduction

Ongoing increase in wide-area network connectivity promise vastly augmented opportunities for collaboration and resource sharing. Now-a-days, various social networking sites like Twitter[1], Facebook[2], MySpace[3], YouTube[4] have gained so much popularity and we cannot ignore them. They have become one of the most important applications of Web 2.0 [1]. They allow people to build connection networks with other people in an easy and timely way and allow them to share various kinds of information and to use a set of services like picture sharing, blogs, wikis etc.

It is evident that the advent of these real-time information networking sites like Twitter have spawned the creation of an unequaled public collection of opinions about every global entity that is of interest. Although Twitter may provision for an excellent channel for opinion creation and presentation, it poses newer and different challenges and the process is incomplete without adept tools for analyzing those opinions to expedite their consumption.

More recently, there have been several research projects that apply sentiment analysis to Twitter corpora in order to extract general public opinion regarding political issues [2]. Due to the increase of hostile and negative communication over social networking sites like Facebook and Twitter, recently the Government of India tried to allay concerns over censorship of these sites where Web users continued to speak out against any proposed restriction on posting of content. As reported in one of the Indian national newspaper [3] "Union Minister for Communications and Information Minister, Kapil Sibal, proposed content screening & censorship of social networks like Twitter and Facebook". Instigated by this the research carried out by us was to use sentiment analysis to gauge the public mood and detect any rising antagonistic or negative feeling on social medias. Although, we firmly believe that censorship is not right path to follow, this recent trend for research for sentiment mining in twitter can be utilized and extended for a gamut of practical applications that range from applications in business (marketing intelligence; product and service bench marking and improvement), applications as sub-component technology (recommender systems; summarization; question answering) to applications in politics. This motivated us to propose a model which retrieves tweets on a certain topic through the Twitter API and calculates the sentiment orientation/score of each tweet.

The area of Sentiment Analysis intends to comprehend these opinions and distribute them into the categories like positive, negative, neutral. Till now most sentiment analysis work has been done on review sites [4]. Review sites provide with the sentiments of products or movies, thus, restricting the domain of application to solely business. Sentiment analysis on Twitter posts is the next step in the field of sentiment analysis, as tweets give us a richer and more varied resource of opinions and sentiments that can be about anything from the latest phone they bought, movie they watched, political issues, religious views or the individuals state of mind. Thus, the foray into Twitter as the corpus allows us to move into different dimensions and diverse applications.

2. Related Work

Applying sentiment analysis on Twitter is the upcoming trend with researchers recognizing the scientific trials and its potential applications. The challenges unique to this problem area are largely attributed to the dominantly

[1]www.twitter.com
[2]www. facebook.com
[3]www.myspace.com
[4]www.youtube.com

informal tone of the micro blogging. Pak and Paroubek [5] rationale the use microblogging and more particularly Twitter as a corpus for sentiment analysis. They cited:

- Microblogging platforms are used by different people to express their opinion about different topics, thus it is a valuable source of people's opinions.
- Twitter contains an enormous number of text posts and it grows every day. The collected corpus can be arbitrarily large.
- Twitter's audience varies from regular users to celebrities, company representatives, politicians, and even country presidents. Therefore, it is possible to collect text posts of users from different social and interests groups.
- Twitter's audience is represented by users from many countries.

Parikh and Movassate [6] implemented two Naive Bayes unigram models, a Naive Bayes bigram model and a Maximum Entropy model to classify tweets. They found that the Naive Bayes classifiers worked much better than the Maximum Entropy model could. Go et al. [7] proposed a solution by using distant supervision, in which their training data consisted of tweets with emoticons. This approach was initially introduced by Read [8]. The emoticons served as noisy labels. They build models using Naive Bayes, MaxEnt and Support Vector Machines (SVM). Their feature space consisted of unigrams, bigrams and POS. The reported that SVM outperformed other models and that unigram were more effective as features. Pak and Paroubek [5] have done similar work but classify the tweets as objective, positive and negative. In order to collect a corpus of objective posts, they retrieved text messages from Twitter accounts of popular newspapers and magazine, such as "New York Times", "Washington Posts" etc. Their classifier is based on the multinomial Naïve Bayes classifier that uses N-gram and POS-tags as features. Barbosa et al. [9] too classified tweets as objective or subjective and then the subjective tweets were classified as positive or negative. The feature space used included features of tweets like retweet, hashtags, link, punctuation and exclamation marks in conjunction with features like prior polarity of words and POS of words.

Mining for entity opinions in Twitter, Batra and Rao[10] used a dataset of tweets spanning two months starting from June 2009. The dataset has roughly 60 million tweets. The entity was extracted using the Stanford NER, user tags and URLs were used to augment the entities found. A corpus of 200,000 product reviews that had been labeled as positive or negative was used to train the model. Using this corpus the model computed the probability that a given unigram or bigram was being used in a positive context and the probability that it was being used in a negative context. Bifet and Frank [11] used Twitter streaming data provided by Firehouse, which gave all messages from every user in real-time. They experimented with three fast incremental methods that were well-suited to deal with data streams: multinomial naive Bayes, stochastic gradient descent, and the Hoeffding tree. They concluded that SGD-based model, used with an appropriate learning rate was the best.

Agarwal et al. [12] approached the task of mining sentiment from twitter, as a 3-way task of classifying sentiment into positive, negative and neutral classes. They experimented with three types of models: unigram model, a feature based model and a tree kernel based model. For the tree kernel based model they designed a new tree representation for tweets. The feature based model that uses 100 features and the unigram model uses over 10,000 features. They concluded features that combine prior polarity of words with their parts-of-speech tags are most important for the classification task. The tree kernel based model outperformed the other two.

The Sentiment Analysis tasks can be done at several levels of granularity, namely, word level, phrase or sentence level, document level and feature level [13]. As Twitter allows its users to share short pieces of information known as "tweets" (limited to 140 characters), the word level granularity aptly suits its setting. Survey through the literature substantiates that the methods of automatically annotating sentiment at the word level fall into the following two categories: (1) dictionary-based approaches and (2) corpus-based approaches. Further, to automate sentiment analysis, different approaches have been applied to predict the sentiments of words, expressions or documents. These include Natural Language Processing (NLP) and Machine Learning (ML) algorithms [14]. In our attempt to mine the sentiment from twitter data we introduce a hybrid approach which combines the advantages of both dictionary & corpus based methods along with the combination of NLP & ML based techniques. The following sections illustrate the proposed paradigm.

3. Data Characteristics

Twitter is a social networking and microblogging service that lets its users post real time messages, called tweets. Tweets have many unique characteristics, which implicates new challenges and shape up the means of carrying sentiment analysis on it as compared to other domains.

Following are some key characteristics of tweets:

- *Message Length:* The maximum length of a Twitter message is 140 characters. This is different from previous sentiment classification research that focused on classifying longer texts, such as product and movie reviews.
- *Writing technique:* The occurrence of incorrect spellings and cyber slang in tweets is more often in comparison with other domains. As the messages are quick and short, people use acronyms, misspell, and use emoticons and other characters that convey special meanings.
- *Availability:* The amount of data available is immense. More people tweet in the public domain as compared to Facebook (as Facebook has many privacy settings) thus making data more readily available. The Twitter API facilitates collection of tweets for training.
- *Topics:* Twitter users post messages about a range of topics unlike other sites which are designed for a specific topic. This differs from a large fraction of past research, which focused on specific domains such as movie reviews.
- *Real time:* Blogs are updated at longer intervals of time as blogs characteristically are longer in nature and writing them takes time. Tweets on the other hand being limited to 140 letters and are updated very often. This gives a more real time feel and represents the first reactions to events.

We now describe some basic terminology related to twitter:

- *Emoticons:* These are pictorial representations of facial expressions using punctuation and letters. The purpose of emoticons is to express the user's mood.
- *Target:* Twitter users make use of the "@" symbol to refer to other users on Twitter. Users are automatically alerted if they have been mentioned in this fashion.
- *Hash tags:* Users use hash tags "#" to mark topics. It is used by Twitter users to make their tweets visible to a greater audience.
- *Special symbols*: "RT" is used to indicate that it is a repeat of someone else's earlier tweet.

4. System Architecture

Opinion words are the words that people use to express their opinion (positive, negative or neutral). To find the semantic orientation of the opinion words in tweets, we propose a novel hybrid approach involving both corpus-based and dictionary-based techniques. We also consider features like emoticons and capitalization as they have recently become a large part of the cyber language.

Fig.1gives the architectural overview of the proposed system.

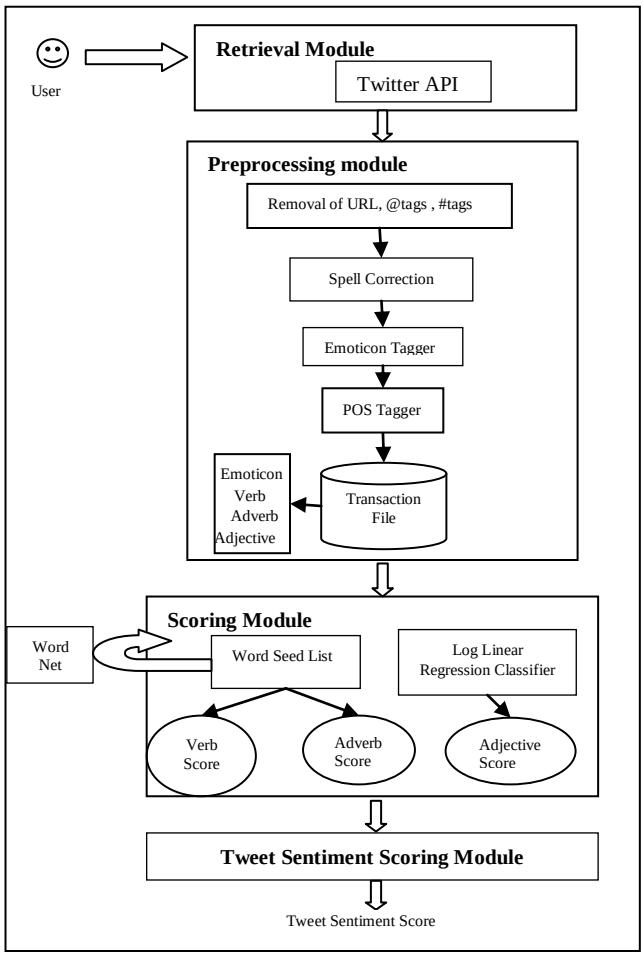

Fig. 1. System Architecture.

To uncover the opinion direction, we will first extract the opinion words in the tweets and then find out their orientation, i.e., to decide whether each opinion word reflects a positive sentiment, negative sentiment or a neutral sentiment. In our work, we are considering the opinion words as the combination of the adjectives along with the verbs and adverbs. The corpus-based method is then used to find the semantic orientation of adjectives and the dictionary-based method is employed to find the semantic orientation of verbs and adverbs. The overall tweet sentiment is then calculated using a linear equation which incorporates emotion intensifiers too.

The following sub-sections expound the details of the proposed system:

4.1 Pre-processing of Tweets

We prepare the transaction file that contains opinion indicators, namely the adjective, adverb and verb along

with emoticons (we have taken a sample set of emoticons and manually assigned opinion strength to them). Also we identify some emotion intensifiers, namely, the percentage of the tweet in Caps, the length of repeated sequences & the number of exclamation marks, amongst others. Thus, we pre-process all the tweets as follows:

a) Remove all URLs (e.g. www.example.com), hash tags (e.g. #topic), targets (@username), special Twitter words ("e.g. RT").
b) Calculate the percentage of the tweet in Caps.
c) Correct spellings; A sequence of repeated characters is tagged by a weight. We do this to differentiate between the regular usage and emphasized usage of a word.
d) Replace all the emoticons with their sentiment polarity (Table 1).
e) Remove all punctuations after counting the number of exclamation marks.
f) Using a POS tagger, the NL Processor linguistic Parser [15], we tag the adjectives, verbs and adverbs.

Table 1: Emoticons

Emoticon	***Meaning***	***Strength***
:D	Big grin	1
BD	Big grin with glasses	1
XD	Laughing	1
\m/	Hi 5	1
:),=),:-)	Happy, smile	0.5
:*	kiss	0.5
:|	Straight face	0
:\	undecided	0
:(	sad	-0.5
</3	Broken heart	-0.5
B(	Sad with glasses	-0.5
:'(	crying	-1
X-(	angry	-1

4.2 Scoring Module

The next step is to find the semantic score of the opinion carriers i.e. the adjectives, verbs and adverbs. As mentioned previously, in our approach we use corpus-based method to find the semantic orientation of adjectives and the dictionary-based method to find the semantic orientation of verbs and adverbs.

4.2.1 Semantic Score of Adjectives

An adjective are a describing word and is used to qualify an object. The semantic orientation of adjectives tend to be domain specific, therefore we use a corpus based approach to quantify the semantic orientation of adjectives in the Twitter domain. Motivated by Hatzivassiloglou and McKeown [16], we ascribe same semantic orientation to conjoined adjectives in most cases and in special cases when the connective is "but", the situation is reversed. Similar to them we apply a log-linear regression model with a linear predictor

$$\eta = w^T x \quad (1)$$

where x is the vector of observed counts in the various conjunction categories(all *and* pairs, all *but* pairs, all attributive *and* pairs, etc.) for the particular adjective pair and w is the vector of weights to be learnt during training. The response y is non-linearly related to η through the inverse logit function

$$y = \frac{e^{\eta}}{1 + e^{\eta}} \quad (2)$$

The value y produced denotes the similarity between the words. The seed list of adjectives was taken and assigned semantic scores manually. We also calculated the semantic score of conjoined adjectives by using the manually assigned scores and the similarity value y.

4.2.2 Semantic Score of Adverbs and Verbs

Although, we can compute the sentiment of a certain texts based on the semantic orientation of the adjectives, but including adverbs is imperative. This is primarily because there are some adverbs in linguistics (such as "not") which are very essential to be taken into consideration as they would completely change the meaning of the adjective which may otherwise have conveyed a positive or a negative orientation.

For example;
One user says, "***This is a good book***" and;
Other says, "***This is not a good book***"

Here, if we had not considered the adverb "not", then both the sentences would have given positive review. On the contrary, first sentence gives the positive review and the second sentence gives the negative review. Further, the strength of the sentiment cannot be measured by merely considering adjectives alone as the opinion words. In other words, an adjective cannot alone convey the intensity of the sentiment with respect to the document in question. Therefore, we take into consideration the adverb strength which modify the adjective; in turn modifying the sentiment strength. Adverb strength helps in assessing whether a document gives a *perfect* positive opinion, *strong* positive opinion, a *slight* positive opinion or a *less* positive opinion.

For example;
One user says, ***"This is a very good book"*** and ;
Other says, ***"This is a good book"***

Some groups of verbs also convey sentiments and opinions (e.g. love, like) and are essential to finding the sentiment strength of the tweet. As adverbs and verbs are not

dependent on the domain, we use dictionary methods to calculate their semantic orientation.

The seed lists of positive and negative adverbs and verbs whose orientation we know is created and then grown by searching in WordNet [17]. Based on intuition, we assign the strengths of a few frequently used adverbs and verbs with values ranging from -1 to +1. We consider some of the most frequently used adverbs and verbs along with their strength as given below in table 2:

Table 2: Verb and Adverb Strengths

Verb	*Strength*	*Adverb*	*Strength*
Love	1	complete	+1
adore	0.9	most	0.9
like	0.8	totally	0.8
enjoy	0.7	extremely	0.7
smile	0.6	too	0.6
impress	0.5	very	0.4
attract	0.4	pretty	0.3
excite	0.3	more	0.2
relax	0.2	much	0.1
reject	-0.2	any	-0.2
disgust	-0.3	quite	-0.3
suffer	-0.4	little	-0.4
dislike	-0.7	less	-0.6
detest	-0.8	not	-0.8
suck	-0.9	never	-0.9
hate	-1	hardly	-1

The complete procedure for predicting adverb and verb polarity is given below:

Procedure "*determine_orientation"* takes the target Adverb/ Verb whose orientation needs to be determined and the respective seed list as the inputs.

1. Procedure **determine_orientation** (target_Adverb/ Verb w_i, Adverb/ Verb_ seedlist)
2. begin
3. if (w_i has synonym s in Adverb/ Verb _ seedlist)
4. { w_i's orientation= *s*'s orientation;
5. add w_i with orientation to Adverb/ Verb _ seedlist ; }
6. else if (w_i has antonym a in Adverb/ Verb _ seedlist)
7. { w_i's orientation = opposite orientation of *a*'s orientation;
8. add w_i with orientation to Adverb/ Verb _ seedlist; }
9. end

The procedure *determine_orientation* searches Word Net and the Adverb/ Verb seed list for each target adjective to predict its orientation (line 3 to line 8). In line 3, it searches synonym set of the target Adverb/ Verb from the Word Net and checks if any synonym has known orientation from the seed list. If so, the target orientation is set to the same orientation as the synonym (line 4) and the target Adverb/ Verb along with the orientation is inserted into the seed list (line 5). Otherwise, the function continues to search antonym set of the target Adverb/ Verb from the Word Net and checks if any Adverb/ Verb have known orientation from the seed list (line 6). If so, the target orientation is set to the opposite of the antonym (line 7) and the target Adverb/ Verb with its orientation is inserted into the seed list (line 8). If neither synonyms nor antonyms of the target word have known orientation, the function just continues the same process for the next Adverb/ Verb since the word's orientation may be found in a later call of the procedure with an updated seed list.

Note:

1) *For those adverbs/ verbs that Word Net cannot recognize, they are discarded as they may not be valid words.*
2) *For those that we cannot find orientations, they will also be removed from the opinion words list and the user will be notified for attention.*
3) *If the user feels that the word is an opinion word and knows its sentiment, he/she can update the seed list.*
4) *For the case that the synonyms/antonyms of an adjective have different known semantic orientations, we use the first found orientation as the orientation for the given adjective.*

4.3 Tweet Sentiment Scoring

As adverbs qualify adjectives and verbs, we group the corresponding adverb and adjective together and call it the adjective group; similarly we group the corresponding verb and adverb together and call it the verb group. The adjective group strength is calculated by the product of adjective score (adj_i) and adverb (adv_i) score, and the verb group strength as the product of verb score (vb_i) and adverb score (adv_i). Sometimes, there is no adverb in the opinion group, so the S (adv) is set as a default value 0.5

To calculate the overall sentiment of the tweet, we average the strength of all opinion indicators like emoticons, exclamation marks, capitalization, word emphasis, adjective group and verb group as shown below:

$$S(T) = \frac{(1 + (P_c + \log(Ns) + \log(N_x))/3)}{|OI(R)|} * \sum_{i=1}^{|OI(R)|} S(AG_i) + S(VG_i) + N_{ei} * S(E_i) \quad (3)$$

Where,
|OI(R)| denotes the size of the set of opinion groups and emoticons extracted from the tweet,
P_c denotes fraction of tweet in caps,
N_s denotes the count of repeated letters,
N_x denotes the count of exclamation marks,
$S(AG_i)$ denotes score of the i^{th} adjective group,
$S(VG_i)$ denotes the score of the i^{th} verb group,
$S(E_i)$ denotes the score of the i^{th} emoticon
N_{ei} denotes the count of the i^{th} emoticon.

P_c, N_s and N_x represent emphasis on the sentiment to be conveyed so they can be collectively called sentiment intensifiers.
If the score of the tweet is more than 1 or less than -1, the score is taken as 1 or -1 respectively.

5. Illustrative Case Study

To clearly illustrate the effectiveness of the proposed method, a case study is presented with a sample tweet:

<tweet>="@kirinv I hate revision, it's BOOOORING!!! I am totally unprepared for my exam tomorrow :(:(Things are not good...#exams"

5.1 The pre-processing of Tweet

A transaction file is created which contains the preprocessed opinion indicators.

5.1.1 Extracting Opinion Intensifiers

The opinion intensifiers are calculated for the tweet as follows.

1) Fraction of tweet in caps:
   There are a total of 18 words in the sentence out of which one is in all caps. Therefore, P_c=1/18=0.055
2) Length of repeated sequence, N_s=3
3) Number of Exclamation marks, N_x=3

5.1.2 Extracting Opinion Words

After the tweet is preprocessed, it is tagged using a POS tagger and the adjective and verb groups are extracted.

The list of Adjective Groups extracted**:**
AG_1=totally unprepared
AG_2=not good
AG_3=boring

The list of Verb Groups extracted:
VG_1=hate

The list of Emoticons extracted:
$E_{1 =}$:(
$N_{e1 =}$ 2

5.2 Scoring Module

Now that we have our adjective group and verb group, we have to find their semantic orientation. Calculation is based on ke

5.2.1 Score of Adjective Group

S (AG_1) = S (totally unprepared) =0.8*-0.5 == -0.4
S (AG_2) = S (not good) =-0.8*1= -0.8
S (AG_3) = S (boring) = 0.5*-0.25 = -0.125

5.2.2 Score of Verb Group

S (VG_1) = S (hate) = 0.5*-0.75 = 0.375

5.3 Tweet Sentiment Scoring

Using the formula defined in equation 3 we can calculate the sentiment strength of the tweet as follows:

$$S(T) = \frac{1.33}{5}\sum_{i=1}^{5} S(AG_i) + S(VG_i) + N_{ei} * S(E_i)$$

$$= \frac{(1.33)}{5} * ((-0.4) + (-0.8) + (-0.125) + (-0.5) + 2*(-0.5))$$
$$= -0.751$$

As we have got a negative value, we can safely classify the tweet as negative.

We applied our approach to a sample set of 10 tweets. The semantic analysis results obtained are depicted in table 3 below.

Table 3: Sample Tweets and semantic orientation

Tweet	***Score***	***Orientation***
@kirinv I hate revision, it's BOOOORING!!! I am totally unprepared for my exam tomorrow :(:(Things are not good...#exams	-0.751	Negative
Criticism of UID laumched yday is extremely unfair. You may hate or even envy Nilekani but can not deny the idea.	0.009	Neutral
"@bigDEElight Keeping it real gone wrong, that was hilarious!! And I wonder how often that actually happens IRL!	0.145	Positive
#iranElection this could get nasty	-0.437	Negative
just getting back from Oaxaca, Mexico by plane	0.125	Positive
I have created a twitter! This is my ONE AND ONLY twitter guys, someone already stole my url. not too happy about it either :(	-0.24	Negative
Happy happy happy :D	0.625	Positive
That was pretty much awesome. :)	0.263	Positive
That other dude sucks!!!	-0.664	Negative
@prncssmojo hey i got a im thingy what is ur screen name?	0	Neutral
Just got home From work. Dam it wuz tough today	-0.281	Negative

The practice result proves that the proposed system has the characteristics of perceiving the semantic orientation of tweets. The results of this work serve as a partial view of the phenomenon. More research needs to be done in order to validate or invalidate these findings, using larger samples.

6. Conclusion

The proliferation of microblogging sites like Twitter offers an unprecedented opportunity to create and employ theories & technologies that search and mine for sentiments. The work presented in this paper specifies a novel approach for sentiment analysis on Twitter data. To uncover the sentiment, we extracted the opinion words (a combination of the adjectives along with the verbs and adverbs) in the tweets. The corpus-based method was used to find the semantic orientation of adjectives and the dictionary-based method to find the semantic orientation of verbs and adverbs. The overall tweet sentiment was then calculated using a linear equation which incorporated emotion intensifiers too. This work is exploratory in nature and the prototype evaluated is a preliminary prototype. The initial results show that it is a motivating technique.

References

[1] L. Colazzo, A. Molinari and N. Villa. "Collaboration vs. Participation: the Role of Virtual Communities in a Web 2.0 world", International Conference on Education Technology and Computer, 2009, pp.321-325.

[2] nlp.stanford.edu/courses/cs224n/2011/reports/patlai.pdf

[3] National Daily, Economic Times: articles.economictimes.indiatimes.com › Collections › Facebook

[4] K. Dave, S. Lawrence, and D.M. Pennock. "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews". In Proceedings of the 12th International Conference on World Wide Web (WWW), 2003, pp. 519–528.

[5] A. Pak and P. Paroubek. "Twitter as a Corpus for Sentiment Analysis and Opinion Mining". In Proceedings of the Seventh Conference on International Language Resources and Evaluation, 2010, pp.1320–1326.

[6] R. Parikh and M. Movassate, "Sentiment Analysis of User-Generated Twitter Updates using Various Classification Techniques", CS224N Final Report, 2009

[7] A. Go, R. Bhayani, L.Huang. "Twitter Sentiment Classification Using Distant Supervision". Stanford University, Technical Paper ,2009

[8] J. Read. "Using emoticons to reduce dependency in machine learning techniques for sentiment classification". In Proceedings of ACL-05, 43nd Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2005

[9] L. Barbosa, J. Feng. "Robust Sentiment Detection on Twitter from Biased and Noisy Data". COLING 2010: Poster Volume, pp. 36-44.

[10] S. Batra and D. Rao, "Entity Based Sentiment Analysis on Twitter", Stanford University,2010

[11] A. Bifet and E. Frank, "Sentiment Knowledge Discovery in Twitter Streaming Data", In Proceedings of the 13th International Conference on Discovery Science, Berlin, Germany: Springer,2010, pp. 1–15.

[12] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, R. Passonneau, "Sentiment Analysis of Twitter Data", In Proceedings of the ACL 2011 Workshop on Languages in Social Media,2011 , pp. 30–38

[13] A. Kumar. and T. M. Sebastian, "Sentiment Analysis: A Perspective on its Past, Present and Future", International Journal of Intelligent Systems and Applications (IJISA), MECS Publisher, 2012 (Accepted to be published)

[14] A. Kumar and T. M. Sebastian, "Machine learning assisted Sentiment Analysis". Proceedings of International Conference on Computer Science & Engineering (ICCSE'2012), 2012, pp. 123-130.

[15] POS Tagger: http://www.infogistics.com/textanalysis.html

[16] V. Hatzivassiloglou and K. McKeown, "Predicting the semantic orientation of adjectives". In Proceedings of the Joint ACL/EACL Conference,2004, pp. 174–181

[17] WordNet: http://wordnet.princeton.edu/

Akshi Kumar is a PhD in Computer Engineering from University of Delhi. She has received her MTech (Master of Technology) and BE (Bachelor of Engineering) degrees in Computer Engineering. She is currently working as a University Assistant Professor in Dept. of Computer Engineering at the Delhi Technological University, Delhi, India. She is editorial review board member for '*The International Journal of Computational Intelligence and Information Security'*, Australia, ISSN: 1837-7823; *'International Journal of Computer Science and Information Security'*, USA, ISSN: 1947-5500; '*Inter-disciplinary Journal of Information, Knowledge & Management'*, published by the Informing Science Institute, USA. (ISSN Print 1555-1229, Online 1555-1237) and '*Webology*', ISSN 1735-188X. She is a life member of Indian Society for Technical Education (ISTE), India, a member of International Association of Computer Science and Information Technology (IACSIT), Singapore, a member of International Association of Engineers (IAENG), Hong Kong, a member of IAENG Society of Computer Science, Hong Kong and a member of Internet Computing Community (ICC), AIRCC. She has many publications to her credit in various journals with high impact factor and international conferences. Her current research interests are in the area of Web Search & Mining, Intelligent Information Retrieval, Web 2.0 & Web Engineering.

Teeja Mary Sebastian is doing M.Tech (Master of Technology) in Computer Technology & Application from Delhi Technological University, Delhi, India and has done her B.Tech (with Distinction) also in Computer Science and Engineering; she is currently working as a scholar in the field of Sentiment Analysis.