Survey of Techniques for Opinion Mining

Nilesh M. Shelke
Assistant Professor, Department of IT, Priyadarshini Indira Gandhi College of Engineering, Nagpur, M.S., India

Shriniwas Deshpande, PhD.
Associate Professor and Head in P. G. Department of Computer Science and Technology, DCPE, HVPM, Amravati. (MS) India.

Vilas Thakre, PhD.
Professor and Head, Post Graduate Department of Computer Science and Engg., SGB Amravati University, Amravati.

ABSTRACT
Opinion mining refers to computational techniques for analyzing the opinions that are extracted from various sources. Existing research work on Opinion is based upon business and e-commerce such as product reviews and movie ratings. Opinion mining involves computational treatment of opinion and subjectivity in text. It has suddenly attracted the attention of the researcher fraternity. This survey paper describes techniques and approaches that promise to directly enable opinion-oriented information seeking systems. An attempt has been made to discuss in de tails various approaches to perform a computational treatment of sentiments and opinions. Various supervised or data-driven techniques for opinion mining like Naïve Byes, Maximum Entropy, SVM are discussed and their strengths and drawbacks are touched upon.

Keywords
Polar expression, opinion mining, POS tagger, entropy, corpus, sentiment, emotion, machine learning.

1. INTRODUCTION
The number of documents expressing opinions is continually increasing on the World Wide Web. People express their views for the mobile products, movies, cars etc. Usually proposed approaches try to find positive or negative opinion features to build training sets and apply classification algorithms (based on several linguistic techniques) to automatically classify new documents extracted from the Web.

Opinion mining can be defined as a sub-discipline of computational linguistics that focuses on extracting people's opinion from the web. It is a *Natural Language Processing* and *Information Extraction* task that aims to obtain writer's feelings expressed in positive or negative comments, questions and requests, by analyzing a large numbers of documents. Generally speaking, sentiment analysis aims to determine the attitude of a speaker or a writer with respect to some topic or the overall tonality of a document. In recent years, the exponential increase in the Internet usage and exchange of public opinion is the driving force behind Opinion Mining today. The Web is a huge repository of structured and unstructured data. The analysis of this data to extract latent public opinion and sentiment is a challenging task.
One can define a sentiment or opinion as a quintuple-
"<o_j, f_{jk}, so_{ijkl}, h_i, t_l >, where o_j is a target object, f_{jk} is a feature of the object o_j, so_{ijkl} is the sentiment value of the opinion of the opinion holder h_i on feature f_{jk} of object o_j at time t_l, so_{ijkl} is +ve,-ve, or neutral, or a more granular rating, hi is an opinion holder, t_l is the time when the opinion is expressed."

The analysis of opinions may be document based where the sentiment in the entire document is summarized as positive, negative or objective. It can be sentence based where individual sentences, bearing sentiments, in the text are classified. Opinion Mining can be phrase based where the phrases in a sentence are classified according to polarity.

Opinion Mining identifies the phrases in a text that bears some sentiment. The author may speak about some *objective facts* or *subjective opinions*. It is necessary to distinguish between the two. Opinion Mining finds the subject towards whom the sentiment is directed. A text may contain many entities but it is necessary to find the entity towards which the sentiment is directed. It identifies the polarity and degree of the sentiment. Sentiments are classified as *objective* (facts), *positive* (denotes a state of happiness, bliss or satisfaction on part of the writer) or *negative* (denotes a state of sorrow, dejection or disappointment on part of the writer). The sentiments can further be given a score based on their *degree* of positivity, negativity or objectivity. The recent expansion of the web encourages users to contribute and express themselves via blogs, videos, social networking sites etc. All these platforms provide a huge amount of valuable information that we are interested to analyse. Given a piece of text, opinion-mining systems analyse:

· Which part is opinion expressing?

· Who wrote the opinion?

· What is being commented?

2. SIGNIFICANCE OF OPINION MINING
Opinion mining applications are the basic infrastructure of large scale collaborative policy-making. They help making sense of thousands of interventions. They help to detect early warning system of possible disruption in a timely manner, by detecting early feedback from citizens. Traditionally, ad hoc surveys are used to collect feedback in a structured manner. However, this kind of data collection is expensive, as it deserves an investment in design and data collection; it is difficult, as people are not interested in answering surveys; and ultimately it is not very valuable, as it detects "known problems" through pre-defined questions and interviewees, but fails to detect the most important problems, the famous "unknown unknown". Opinion mining is helpful to identify problems by listening, rather than by asking, thereby ensuring a more accurate reflection of reality.

Argument mapping software is then useful to ensure that policy debates are logical and evidence-based, and do not repeat the same arguments again and again.

These tools would finally be helpful not only for policy-makers, but also for citizens who could more easily understand the key points of a discussion and participate to the policy-making process.

3. RELATED WORK

Many classifiers for opinion mining have been recommended in this decade. Among them naive-bayes, support vector machines, decision tress, rule-based classifiers are important and widely used in several applications. Those are discussed separately in section 4.
We have categorized the work done for classification of sentiment analysis as follows:

3.1 Word or Phrase Sentiment Classification

Word Sentiment Classification has become the basis for further phrase and document classification. Handcrafted and semi-handcrafted dictionaries were built up [1,2].

The words in them were mostly adjectives or adverbs that have semantic orientation [3, 4, 5] and the orientation was defined by researchers. The approaches to classify sentiment at word level could be grouped into two: 1) Corpus based approaches and 2) Dictionary based approaches. This has been discussed in section 4.

3.2 Document Sentiment Classification

Supervised machine learning approaches are popular among researchers in predicting the overall sentiment of the document [6, 7, 8]. The objective of this task is to classify each review document as expressing a positive or a negative sentiment about an object (e.g., a movie, a camera, or a car). Most of them focused on labeling a new sample as “positive” or “negative” based on previously seen samples annotated by humans. Grading a review on a multi-variant scale is fairly a new application in this area. The entire process is typically composed of two steps:

1) Extracting the subjective features from the training data and converting them as feature vectors.

2) Training the classifier on the feature vectors and apply the classification on a new sample. Preprocessing the raw documents before extracting the subjective features is also done. The preprocessing stage includes removing HTML tags, tokenization of documents.

3.3 Sentence Level Sentiment Classification

Since, researchers thought it was too coarse to compute the sentiment at document level, they investigated approaches to determine the focus of each sentence. Several researchers also studied sentence-level sentiment classification [9, 10, 11] i.e., classifying each sentence as expressing a positive or a negative opinion. The model of feature-based opinion mining and summarization is proposed in [12, 13]. They computed the semantic orientation at sentence level. They extracted opinion bearing terms, opinion holders and opinion-product aspect association in each sentence and then analysed the semantic orientation. There was also an area of research called *aspect based sentiment classification* where they extract aspects of a product and rate sentiments of people on its each aspect. Earlier researchers conducted experiments on aspect based classification of movie reviews. They used information extraction techniques like pronoun resolution, entity extraction and con-referencing to segment each sentence. They predicted the sentiment of users towards cast (producers, directors) and also overall sentiment.

4. SENTIMENT CLASSIFICATION

The literature survey done indicates two types of techniques include machine learning and semantic orientation.

4.1. Machine Learning

Several Machine Learning methods have been studied. Prominent methods are: Naive Bayes Classification, Maximum Entropy Classification, and Support Vector Machines.

In his work, Pang Lee *et al.* [14, 15], compared the performance of Naïve Bayes, Maximum Entropy and Support Vector Machines in SA on different features like considering only unigrams, bigrams, combination of both, incorporating parts of speech and position information, taking only adjectives *etc*. The result has been summarized in the Table 1. It is observed from the results that:

- Feature presence is more important than feature frequency.
- Using Bigrams the accuracy actually falls.
- Accuracy improves if all the frequently occurring words from all parts of speech are taken, not only Adjectives.
- Incorporating position information increases accuracy.
- When the feature space is small, Naïve Bayes performs better than SVM. But SVM’s perform better when feature space is increased.

When feature space is increased, Maximum Entropy may perform better than Naïve Bayes but it may also suffer from overfitting. Despite its simplicity and the fact that its conditional independence assumption clearly does not hold in real-world situations, Naive Bayes-based text categorization still tends to perform surprisingly well [16]; indeed, Domingos and Pazzani [17] show that Naive Bayes is optimal for certain problem classes with highly dependent features. Maximum Entropy Classification (Max Ent, or ME, for short) is an alternative technique which has proven effective in a number of natural language processing applications [18]. Nigam et al. [19,20] show that it sometimes, but not always, outperforms Naive Bayes at standard text classification.

Table 1. Comparison of Accuracies using Naive Bayes, Maximum Entropy and Support Vector Machines

Features	No. of features	Frequency or Presence	NB	ME	SVM
Unigrams	16165	Freq.	78.7	N/A	72.8
Unigrams	16165	Pres.	81.0	80.4	82.9
Unigrams+Bigrams	32330	Pres.	80.6	80.8	82.7
Bigrams	16165	Pres.	77.3	77.4	77.1
Unigrams+POS	16695	Pres.	81.5	80.4	81.9
Adjectives	2633	Pres.	77.0	77.7	75.1
Total 2633 unigrams	2633	Pres.	80.3	81.0	81.4
Unigrams+position	22430	Pres.	81.0	80.1	81.6

4.2 Sementic Orientation

Problem of Opinion mining can be categorized as sentiment classification and feature based opinion mining.

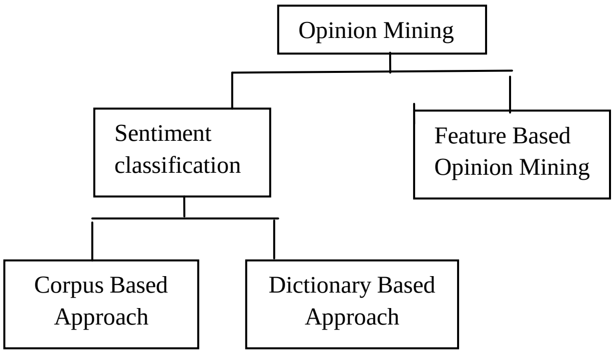

Fig 1: Classification of Approaches of Sentiment Orientation.

Problem of Opinion mining can be categorized as sentiment classification and feature based opinion mining.

Corpus Based Approach: Popular corpus-driven method is to determine the emotional affinity of words which is to learn their probabilistic affective scores from large corpora. Mihalcea and Liu [21] have used this method to assign a *happiness factor* to words depending on the frequency of their occurrences in happy-labeled blogposts compared to their total frequency in a corpus containing blogposts labeled with "happy" and "sad" mood annotations. They also compare the happiness factor scores of words with the scores in the ANEW list . The ANEW list, prepared on the basis of psychological experiments, assigns scores to words along the three dimensions of affect in the PAD model [22]. These dimensions are: *Pleasure/displeasure*, *Arousal/non-arousal*, and *Dominance/submissiveness*.

Dictionary Based Approach: These approaches have used lexical resources such as WordNet to automatically acquire emotion-related words for emotion classification experiments. Starting from a set of primary emotion adjectives, Alm et al. [23] retrieve similar words from WordNet utilizing all senses of all words in the synsets that contain the emotion adjectives.

Researchers also exploit the synonym and hyponym relations in WordNet to manually find words similar to nominal emotion words. Strapparava et al. [24] have used WordNet-Affect, an affective extension of WordNet in their experiments to automatically detect emotion in text. They use five of the six basic emotional categories described by Ekman [25]. For directly affective words, they use weights from WordNet-Affect. However, for indirectly affective words, their approach is to assign them affective weights based on their semantic similarity to an emotional category. The affective weights are automatically acquired from a very large text corpus in an unsupervised fashion. The approach of using sentiment orientation of constituting words to determine the overall sentiment of the document suffers from drawbacks, as it relies only on superficial features, whereas sentiment is often communicated through the composite meaning of the text, rather than exclusively through the use of affect words.

The problem of extracting the semantic orientation (SO) of a text (i.e., whether the text is positive or negative towards a particular subject matter) often takes as a starting point the problem of determining semantic orientation for individual words. The hypothesis is that, given the SO of relevant words in a text, we can determine the SO for the entire text. However, if we assume that SO for individual words is an important part of the problem, then we need lists of words with their corresponding SO, since such information is not typically contained in a traditional dictionary.

Whitelaw et al. [26] use a semi-automatic method to create a dictionary of words that express appraisal. Appraisal is a functional framework for describing evaluation in text: how personal feelings, judgment about other people, and appreciation of objects and art are expressed. Word similarity may be another way of building dictionaries, starting from words whose SO we already know. Manual and semi-automatic methods, although highly accurate, are not ideal, given that it is time-consuming and labour-intensive to compile a list of all the words that can possibly express sentiment.

The Semantic orientation approach to Sentiment analysis is unsupervised learning because it does not require prior training in order to mine the data. Instead, it measures how far a word is inclined towards positive and negative.

Earlier research carried out for unsupervised sentiment classification makes use of lexical resources available. Kamps et al [27] focused on the use of lexical relations in sentiment classification. Andrea Esuli and Fabrizio Sebastiani [28] proposed semi-supervised learning method started from expanding an initial seed set using WordNet. Their basic assumption is terms with similar orientation tend to have similar glosses. They determined the expanded seed terms semantic orientation through gloss classification by statistical technique.

When the review where an opinion lies in, cannot provide enough contextual information to determine the orientation of opinion, Chunxu Wu [29] proposed an approach which resort to other reviews discussing the same topic to mine useful contextual information, then use semantic similarity measures to judge the orientation of opinion. They attempted to tackle this problem by getting the orientation of context independent opinion, then consider the context dependent opinions using linguistic rules to infer orientation of context distinct-dependent opinion, then extract contextual information from other reviews that comment on the same product feature to judge the context indistinct-dependent opinions.

An unsupervised learning algorithm by extracting the sentiment phrases of each review by rules of part-of-speech (POS) patterns was investigated by Ting-Chun Peng and Chia-Chun Shih [30]. For each unknown sentiment phrase, they used it as a query term to get top-N relevant snippets from a search engine respectively. Next, by using a gathered sentiment lexicon, predictive sentiments of unknown sentiment phrases are computed based on the sentiments of nearby known sentiment words inside the snippets. They consider only opinionated sentences containing at least one detected sentiment phrase for opinion extraction. Using the POS pattern opinion extraction is done.

Gang Li & Fei Liu [31] developed an approach based on the k-means clustering algorithm. The technique of TF-IDF (term frequency – inverse document frequency) weighting is applied on the raw data. Then, a voting mechanism is used to extract a more stable clustering result. The result is obtained based on multiple implementations of the clustering process. Finally, the term score can be used to further enhance the clustering result. Documents are clustered into positive group and negative group. Chaovalit and Zhou [32] compared the Semantic Orientation approach with the N-gram model machine learning approach by applying to movie reviews.

They confirmed from the results that the machine learning approach is more accurate but requires a significant amount of time to train the model. In comparison, the semantic orientation approach is slightly less accurate but is more efficient to use in real-time applications. The performance of semantic orientation altso relies on the performance of the underlying POS tagger.

4.2.1 Features for Opinion Mining

Feature engineering is an extremely basic and essential task for Opinion Mining. Converting a piece of text to a feature vector is the basic step in any data driven approach to Opnion. Some commonly used features in Opinion Mining and their critiques have been discussed in the following sections.

Term Presence vs. Term Frequency

Term frequency has always been considered essential in traditional Information Retrieval and Text Classification tasks. But it is found that *term presence* is more important to Sentiment analysis than *term frequency*. That is, binary-valued feature vectors in which the entries merely indicate whether a term occurs (value 1) or not (value 0). It has also been seen that the occurrence of rare words contain more information than frequently occurring words, a phenomenon called *Hapax Legomena*.

Term Position

Words appearing in certain positions in the text carry more sentiment or weightage than words appearing elsewhere. This is similar to IR where words appearing in topic Titles, Subtitles or Abstracts *etc* are given more weightage than those appearing in the body. In many examples, although the text contains positive words throughout, the presence of a negative sentiment at the end sentence plays the deciding role in determining the sentiment. Thus generally words appearing in the 1st few sentences and last few sentences in a text are given more weightage than those appearing elsewhere.

N-gram Features

N-grams are capable of capturing context to some extent and are widely used in Natural Language Processing tasks. Whether higher order n-grams are useful is a matter of debate. It has been reported by researchers that unigrams outperform bigrams when classifying movie reviews by sentiment polarity, but other researchers found that in some settings, bigrams and trigrams perform better.

Parts of Speech

Parts of Speech information is most commonly exploited in all NLP tasks. One of the most important reasons is that they provide a crude form of word sense disambiguation.

Adjectives only

Adjectives have been used most frequently as features amongst all parts of speech. A strong correlation between adjectives and subjectivity has been found. Although all the parts of speech are important people most commonly used adjectives to depict most of the sentiments and a high accuracy have been reported by all the works concentrating on only adjectives for feature generation. Pang Lee *et al.* [15] achieved as accuracy of around 82.8% in movie review domains using only adjectives in movie review domains.

Table 2. Word List containing Positive and Negative Adjectives

	Proposed word list
Human 1	positive: dazzling, brilliant, phenomenal, excellent, fantastic. negative: suck, terrible, awful, unwatchable, hideous
Human 2	positive: gripping, mesmerizing, riveting, spectacular, cool, awesome, thrilling, badass, excellent, moving, exciting, negative: bad, clichéd, sucks, boring, stupid, slow.

Adjective-Adverb Combination

Most of the adverbs have no prior polarity. But when they occur with sentiment bearing adjectives, they can play a major role in determining the sentiment of a sentence. Adverbs alter the sentiment value of the adjective that they are used with. Adverbs of degree, on the basis of the extent to which they modify this sentiment value, are classified as:

- Adverbs of affirmation: certainly, totally
- Adverbs of doubt: maybe, probably
- Strongly intensifying adverbs: exceedingly, immensely
- Weakly intensifying adverbs: barely, slightly
- Negation and minimizers: never

5. KEY APPLICATIONS

Opinions are so important that whenever one needs to make a decision, one wants to hear others' opinions. This is true for both individuals and organizations. The technology of opinion mining thus has a tremendous scope for practical applications.

Individual consumers: If an individual wants to purchase a product, it is useful to see a summary of opinions of existing users so that he/she can make an informed decision. This is better than reading a large number of reviews to form a mental picture of the strengths and weaknesses of the product. He/she can also compare the summaries of opinions of competing products, which is even more useful.

Organizations and businesses: Opinion mining is equally, if not even more, important to businesses and organizations. For example, it is critical for a product manufacturer to know how consumers perceive its products and those of its competitors. This information is not only useful for marketing and product benchmarking but also useful for product design and product developments.

Sentiment detection has a wide variety of applications in information systems, including classifying reviews, summarizing review and other real time applications. There are likely to be many other applications that is not discussed. It is found that sentiment classifiers are severely dependent on domains or topics. From the above work it is evident that neither classification model consistently outperforms the other, different types of features have distinct distributions. It is also found that different types of features and classification algorithms are combined in an efficient way in order to overcome their individual drawbacks and benefit from each other merits, and finally enhance the sentiment classification performance. In future, more work is needed on further improving the performance measures. Sentiment analysis can be applied for new applications.

Although the techniques and algorithms used for sentiment analysis are advancing fast, however, a lot of problems in this field of study remain unsolved. The main challenging aspects exist in use of other languages, dealing with negation expressions; produce a summary of opinions based on product features/attributes, complexity of sentence/ document, handling of implicit product features, etc. More future research could be dedicated to these challenges.

6. CONCLUSION

In this paper various approaches to Sentiment Analysis, mainly Machine Learning have been studdied. It also provides a detailed view of the different applications of Sentiment Analysis. Working on a corpus of blog sentences annotated with emotion labels, it can be surveyed that a combination of corpus-based unigram features and features derived from emotion lexicons can help automatically distinguish basic emotion categories in written text. In addition, a method has been described that builds an emotion lexicon derived from Thesaurus on the basis of semantic relatedness of words to a set of basic emotion words for each emotion category. The effectiveness of this emotion lexicon can be demonstrated in the emotion classification tasks.

Feature engineering, as in several Machine Learning and Natural Language Processing applications, plays a vital role in Opnion Mining. We have seen the use of phrases as well as words as features. It has been seen that Adjectives as word features can capture majority of the sentiment.

7. REFERENCES

[1] V. Hatzivassiloglou and K.R. McKeown. Predicting the semantic orientation of adjectives. In *proceedings of 35th ACL.*, 1997.

[2] V. Hatzivassiloglou and J. Wiebe. Effects of adjective orientation and gradability on sentence subjectivity. In *proceedings of 18th International Conference on Computational Linguistics.*, 2000.

[3] A. Andreevskaia and S. Bergler. “Mining wordnet for a fuzzy sentiment: Sentiment Tag Extraction from Wordnet Glosses”. In *proceedings of EACL.*, 2006.

[4] A. Esilu and F. Sebastini. Sentiwordnet: it is publicly available resource for opinion mining. *In proceedings of LREC 2006,* 2006.

[5] M. Gamon and A. Aue. Automatic Identification of Sentiment Vacabulary Exploiting Low Association with known sentiment terms. *In proceedings of ACL workshop on feature engineering in machine learning in NLP, 2005.*

[6] Dave, D., Lawrence, A., and Pennock, D. “Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews”. *Proceedings of International World Wide Web Conference* (*WWW’03*), 2003.

[7] Pang, B., Lee, L. and Vaithyanathan, S. “Thumbs up? Sentiment Classification Using Machine Learning Techniques”. *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing* (*EMNLP’02*), 2002.

[8] Turney, P. “Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews”. *ACL’02,* 2002.

[9] Kim, S. and Hovy, E. “Determining the Sentiment of Opinions”. *Proceedings of the 20th International Conference on Computational Linguistics* (*COLING’04*), 2004.

[10] Wiebe, J. and Riloff, E. “Creating Subjective and Objective Sentence Classifiers from Unannotated Texts”. *Proceedings of International Conference on Intelligent Text Processing and Computational Linguistics* (*CICLing’05*), 2005.

[11] Wilson, T., Wiebe, J. and Hwa, R., “Just How Mad Are You? Finding Strong and Weak Opinion Clauses”. *Proceedings of National Conference on Artificial Intelligence* (*AAAI’04*), 2004.

[12] Hu, M and Liu, B. “Mining and Summarizing Customer Reviews”. *Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (*KDD’04*), 2004.

[13] Liu, B., Hu, M. and Cheng, J. “Opinion Observer: Analyzing and Comparing Opinions on the Web”, *Proceedings of International World Wide Web Conference* (*WWW’05*), 2005.

[14] Pang, Bo and Lee, Lillian and Vaithyanathan, Shivakumar, Thumbs up? Sentiment Classification using Machine Learning Techniques, *Proceedings of EMNLP 2002.*

[15] Pang, Bo and Lee, Lillian and Vaithyanathan, Shivakumar, Thumbs up?: Sentiment Classification using machine learning techniques, *In Proceedings of the ACL-02 conference on Empirical Methods in Natural Language, 2002*

[16] David D. Lewis. 1998. “The Independence Assumption in Information Retrieval”. In *Proc. of the European Conference on Machine Learning (ECML)*, p.p. 4–15.

[17] Pedro Domingos and Michael J. Pazzani. 1997, “On the Optimality of the Simple Bayesian Classifier Under Zero-One Loss. *Machine Learning”*, 29(2-3):103–130.

[18] Adam L. Berger, Stephen A. Della Pietra, and Vincent J. Della Pietra. 1996. “A Maximum Entropy Approach to Natural Language Processing”. *Computational Linguistics*, 22(1):39–71.

[19] Andrew McCallum and Kamal Nigam. 1998. “A Comparison of Event Models for Naive Bayes Text Classification”. In *Proc. of the AAAI-98 Workshop on Learning for Text Categorization*, pages 41–48.

[20] Stanley Chen and Ronald Rosenfeld. 2000. “A Survey of Smoothing Techniques for ME Models”. *IEEE Trans. Speech and Audio Processing,* 8(1):37–50.

[21] Mihalcea, R. and Liu, H. (2006). “A Corpus-Based Approach to Finding Happiness”, in the *AAAI Spring Symposium on Computational Approaches to Weblogs*, Stanford, California, USA.

[22] Mehrabian, A. (1995). “Framework for a Comprehensive Description and Measurement of Emotional States”. *Genetic, Social, and General Psychology Monographs,* 121, p.p. 339-361.

[23] Alm, C.O., Roth, D. and Sproat, R. (2005). Emotions from text: machine learning for text based emotion prediction. In *Proceedings of the Joint Conference on Human Language Technology / Empirical Methods in*

Natural Language Processing (HLT/EMNLP 2005),Vancouver, Canada, p.p. 579-586.

[24] Mihalcea, R. and Strapparava, C. (2005). Making Computers Laugh: Investigations in Automatic Humor Recognition, In *Proceedings of the Joint Conference on Human Language Technology/Empirical Methods in Natural Language Processing(HLT/EMNLP)*, p.p. 531-538, Vancouver, Canada.

[25] Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion*, p.p. 169-200.

[26] Whitelaw, C., Garg, N. & Argamon, S. (2005). Using Appraisal groups for sentiment analysis. Proceedings of ACM SIGIR Conference on Information and Knowledge Management (CIKM 2005) p. p. 625 -631.

[27] 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.

[28] Andrea Esuli and Fabrizio Sebastiani, "Determining the Semantic Orientation of Terms through Gloss Classification", Proceedings of 14th ACM International Conference on Information and Knowledge Management, p.p. 617-624, Bremen, Germany, 2005.

[29] Chunxu Wu, Lingfeng Shen, "A New Method of Using Contextual Information to Infer the Semantic Orientations of Context Dependent Opinions", 2009 International Conference on Artificial Intelligence and Computational Intelligence.

[30] Ting-Chun Peng and Chia-Chun Shih , "An Unsupervised Snippet-based Sentiment Classification Method for Chinese Unknown Phrases without using Reference Word Pairs", 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology Journal Of Computing, Volume 2 (8), August 2010, ISSN, p.p. 2151-9617 .

[31] Gang Li, Fei Liu, "A Clustering-based Approach on Sentiment Analysis", 2010, 978-1-4244-6793-8/10 ©2010 IEEE.

[32] Chaovalit, Lina Zhou, "Movie Review Mining: a Comparison between Supervised and Unsupervised Classification Approaches", Proceedings of the 38th Hawaii International Conference on System Sciences – 2005.