I.J. Intelligent Systems and Applications, 2012, 10, 1-14
Published Online September 2012 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijisa.2012.10.01

Sentiment Analysis: A Perspective on its Past, Present and Future

Akshi Kumar
Department of Computer Engineering, Delhi Technological University, Delhi, India
Email:akshi.kumar@gmail.com

Teeja Mary Sebastian
Department of Computer Engineering, Delhi Technological University, Delhi, India
Email: teeja.sebastian@gmail.com

***Abstract*—**The proliferation of Web-enabled devices, including desktops, laptops, tablets, and mobile phones, enables people to communicate, participate and collaborate with each other in various Web communities, viz., forums, social networks, blogs. Simultaneously, the enormous amount of heterogeneous data that is generated by the users of these communities, offers an unprecedented opportunity to create and employ theories & technologies that search and retrieve relevant data from the huge quantity of information available and mine for opinions thereafter. Consequently, Sentiment Analysis which automatically extracts and analyses the subjectivities and sentiments (or polarities) in written text has emerged as an active area of research. This paper previews and reviews the substantial research on the subject of sentiment analysis, expounding its basic terminology, tasks and granularity levels. It further gives an overview of the state- of – art depicting some previous attempts to study sentiment analysis. Its practical and potential applications are also discussed, followed by the issues and challenges that will keep the field dynamic and lively for years to come.

***Index Terms*—** Sentiment Analysis, Opinion, Web 2.0, Tasks, Levels, Applications, Issues

I. Introduction

A vital part of the information era has been to find out the opinions of other people. In the pre-web era, it was customary for an individual to ask his or her friends and relatives for opinions before making a decision. Organizations conducted opinion polls, surveys to understand the sentiment and opinion of the general public towards its products or services. In the past few years, web documents are receiving great attention as a new medium that describes individual experiences and opinions. With proliferation of Web 2.0 [1] applications such as micro-blogging, forums and social networks came. Reviews, comments, recommendations, ratings, feedbacks were generated by users. Hence, with the advent of World Wide Web[1] and specifically with the growth and popularity of Web 2.0 where focus shifted to user generated content, the way people express opinion or their view has changed dramatically. People can now make their opinion, views, sentiment known on their personal websites, blogs, social networking sites, forums and review sites. They are comfortable with going online to get advice. Organizations have evolved and now look at review sites to know how the public has received their product instead of conducting surveys. This information available on the Web is a valuable resource for marketing intelligence, social psychologists and others interested in extracting and mining views, moods and attitude [2].

There is a vast amount of information available on the Web which can assist individuals and organization in decision making processes but at the same time present many challenges as organizations and individuals attempt to analyze and comprehend the collective opinion of others. Unfortunately finding opinion sources, monitoring them and then analyzing them are herculean tasks. It is not possible to manually find opinion sources online, extract sentiments from them and then to express them in a standard format. Thus the need to automate this process arises and sentiment analysis [3] is the answer to this need.

Sentiment analysis or Opinion mining, as it is sometimes called, is one of many areas of computational studies that deal with opinion oriented natural language processing. Such opinion oriented studies include among others, genre distinctions, emotion and mood recognition, ranking, relevance computations, perspectives in text, text source identification and opinion oriented summarization [4]. Sentiment analysis has turned out as an exciting new trend in social media with 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;

[1]World Wide Web Consortium(W3C) www.w3.org/WWW/

Copyright © 2012 MECS

summarization; question answering) to applications in politics. It has great potential to be used in business strategies and has helped organizations get a real-time feedback loop about their marketing strategy or advertisements from the reaction of the public through tweets, posts and blogs. For a new product launch it can give them instant feedback about the reception of the new product. It can gauge what their brand image is, whether they are liked or not.

As the field of sentiment analysis is relatively new, the terminology used to describe this field of research is many. The terms opinion mining, subjectivity analysis, review mining and appraisal extraction are used interchangeably with sentiment analysis. Subjectivity analysis or subjectivity classification is focused on the task of whether the sentence or document is expressing opinions or sentiments of the author or just merely stating facts. Majority of the papers which use the phrase "sentiment analysis" focus on the specific application of classifying reviews as to their polarity (either positive or negative) [4]. The term opinion mining was first noticed in a paper by Dave et al. [5]. The paper defined that an opinion mining tool would "process a set of search results for a given item, generating a list of product attributes (quality, features, etc.) and aggregating opinions about each of them (poor, mixed, good)". This definition has been broadened to include various other works in this area. The evolution of the phrase sentiment analysis is similar to that of Opinion Mining. We have used these terms interchangeably in this paper.

Recently a lot of interest has been generated in the field of sentiment analysis, with researchers recognizing the scientific trials and potential applications supported by the processing of subjective language. Some factors substantiated by research till date, that push the development of the research area, include, augmenting of machine learning methods in natural language processing and information retrieval, increase in World Wide Web to provide training datasets for machine learning algorithms and the realization of commercial and intelligent applications that the area provides. As an example of one of the latest applications of sentiment analysis, Twitter[1], Inc. incorporated an advanced tweet-searching function based on sentiment direction, where users can search for positive or negative tweets on a particular topic.

This paper gives an overview of sentiment analysis, its basic terminology, tasks and levels and discusses practical and potential applications of sentiment analysis further expounding its significant open research directions. The paper is organized as follows: the first section introduces sentiment analysis and discusses its history. It is followed by a section which explains the basic terminology. Section 3 expounds how different Web 2.0 applications add dimensions to the sentiment analysis tasks, which are illustrated in section 4 followed by section 5 which explains the granularity at which these tasks can be performed. Section 6 explicates the current state- of- art and describes how machine learning has proved its worth as a technique used for solving the sentiment analysis tasks. Section 7 presents the various applications of sentiment analysis. Lastly, section 8 discusses the various issues that turn out as open problems to be addressed which urge researchers to make significant improvements to understand and work in the sentiment analysis domain.

II. Basic Terminology of Sentiment Analysis

Formally stating Sentiment Analysis is the computational study of opinions, sentiments and emotions expressed in text [3]. The goal of sentiment analysis is to detect subjective information contained in various sources and determine the mind-set of an author towards an issue or the overall disposition of a document.

Wiebe et al. [6] described subjectivity as the linguistic expression of somebody's opinions, sentiments, emotions, evaluations, beliefs and speculations. The words opinion, sentiment, view and belief are used interchangeably but there are subtle differences between them [4].

- *Opinion:* A conclusion thought out yet open to dispute ("each expert seemed to have a different opinion").
- *View:* subjective opinion ("very assertive in stating his views").
- *Belief:* deliberate acceptance and intellectual assent ("a firm belief in her party's platform").
- *Sentiment:* a settled opinion reflective of one's feelings ("her feminist sentiments are well-known").

Sentiment analysis is done on user generated content on the Web which contains opinions, sentiments or views. An opinionated document can be a product review, a forum post, a blog or a tweet, that evaluates an object. The opinions indicated can be about anything or anybody, for e.g. products, issues, people, organizations or a service.

Lui [3] mathematically represented an opinion as a quintuple (o, f, so, h, t), where o is an object; f is a feature of the object o; so is the orientation or polarity of the opinion on feature f of object o; h is an opinion holder; t is the time when the opinion is expressed.

- *Object:* An entity which can be a product, person, event, organization, or topic. The object can have attributes, features or components associated with it. Further on the components can have subcomponents and attributes

[1] http://twitter.com/

- *Feature:* An attribute (or a part) of the object with respect to which evaluation is made.
- *Opinion orientation or polarity:* The orientation of an opinion on a feature f indicates whether the opinion is positive, negative or neutral. Most work has been done on binary classification i.e. into positive or negative. But opinions can vary in intensity from very strong to weak [7]. For example a positive sentiment can range from content to happy to ecstatic. Thus, strength of opinion can be scaled and depending on the application the number of levels can be decided.
- *Opinion holder:* The holder of an opinion is the person or organization that expresses the opinion.

The following example in Fig. 1 illustrates the basic terminology of sentiment analysis:

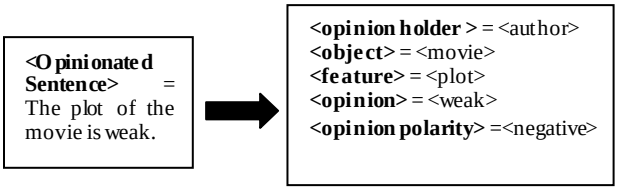

Fig.1 Example corresponding to Terminology of Sentiment Analysis

III. Web 2.0 and Sentiment Analysis

The term Web 2.0 was made popular by Tim O'Rielly in the O'Reilly Media Web 2.0 conference in late 2004. Web 2.0 is an evolution from passive viewing of information to interactive creation of user generated data by the collaboration of users on the Web. Every facet of Web 2.0 is driven by contribution and collaboration. The evolution of Web from Web 1.0 to Web 2.0 was enabled by the rise of read/write platforms such as blogging, social networks, and free image and video sharing sites. These platforms have jointly allowed exceptionally effortless content creation and sharing by anyone.

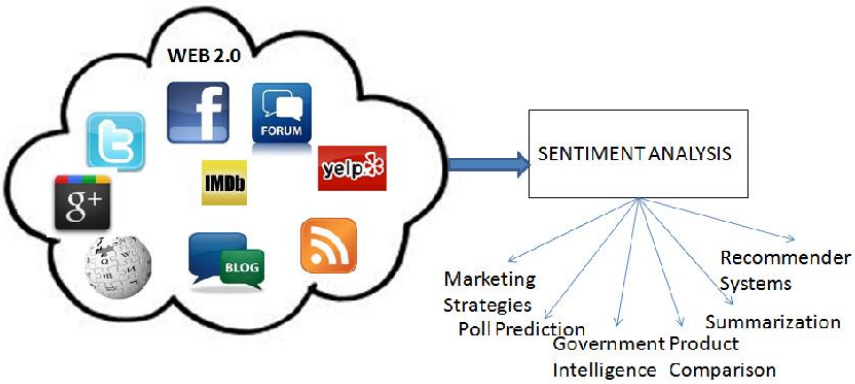

Fig.2 Conceptual model of Sentiment Analysis

The research field of sentiment analysis has been rapidly progressing because of the rich and diverse data provided by Web 2.0 applications. Blogs, review sites, forums, microblogging sites, wikis and social networks have all provided different dimensions to the data used for sentiment analysis.

A. *Review Sites*

A review site is a website which allows users to post reviews which give a critical opinion about people, businesses, products, or services. Most sentiment analysis work has been done on movie and product review sites [5, 7, 9]. The purpose of a review is to appraise a specific object, thus it is a single domain problem. Sentiment analysis on review sites is useful to both manufacturers and potential consumers of the product. The manufacturers can gauge the reception of a product based on the reviews. They can derive the features liked and disliked by the reviewers.

B. *Blogs*

The term web-log or blog, refers to a simple webpage consisting of brief paragraphs of opinion, information, personal diary entries, or links, called posts, arranged chronologically with the most recent first, in the style of an online journal [10]. The bloggers post at hourly, daily or weekly basis which makes the interactions faster and more real-time. Different blogs have different styles of presentation, content material and writing techniques. Sentiment analysis on blogs [11, 12, 13] has been used to predict movie sales, political mood and sales analysis.

C. *Forums*

Forums or message boards allow its members to hold conversations by posting on the site. Forums are generally dedicated to a topic and thus using forums as a database allows us to do sentiment analysis in a single domain.

D. *Social Networks*

Social networking is online services or sites which try to emulate social relationships amongst people who know each other or share a common interest. Social networking sites allow users to share ideas, activities, events, and interests within their individual networks.

Copyright © 2012 MECS

Social network posts can be about anything from the latest phone bought, movie watched, political issues or the individual's state of mind. Thus posts give us a richer and more varied resource of opinions and sentiments.

1) *Twitter*

Twitter is an online social networking and micro blogging service that enables its users to send and read text-based posts of up to 140 characters, known as "tweets". Sentiment analysis on twitter [14, 15, 16] is an upcoming trend with it being used to predict poll results [17] among various other applications.

2) *Facebook*

Facebook[1] is a social networking service and website launched in February 2004. The site allows users to create profiles for themselves, upload photographs and videos. Users can view the profiles of other users who are added as their friends and exchange text messages.

Social media is the new source of information on the Web. It connects the entire world and thus people can much more easily influence each other. The remarkable increase in the magnitude of information available calls for an automated approach to respond to shifts in sentiment and rising trends

IV. Sentiment Analysis Tasks

Sentiment analysis is a challenging interdisciplinary task which includes natural language processing, web mining and machine learning. It is a complex task and encompasses several separate tasks, viz:

- Subjectivity Classification
- Sentiment Classification
- Complimentary Tasks
  - Object Holder Extraction
  - Object/ Feature Extraction

Fig. 3 illustrates the major tasks in a sentiment analysis:

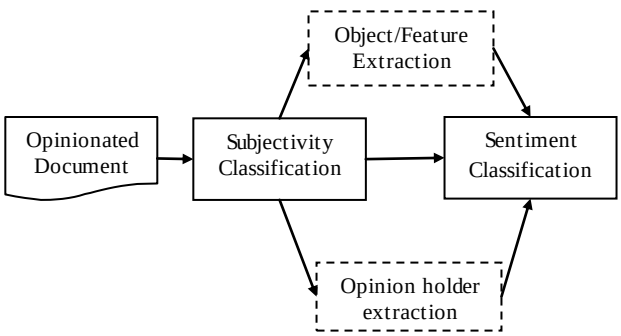

Fig.3 Tasks of Sentiment Analysis

[1] https://www.facebook.com/

The following subsections expound the details of the major tasks in Sentiment Analysis:

A. *Subjectivity classification*

Typically, any given document will contain sentences that express opinion and some that do not. That is, a document is a collection of objective sentences, sentences that state a fact, and subjective sentences, sentences that represents the author's opinion, point of view or emotion. Subjectivity classification is the task of classifying sentences as opinionated or not opinionated [18, 19]. Tang et al. [2], stated subjectivity classification as follows: Let $S = \{s_1, \ldots, s_n\}$ be a set of sentences in document D. The problem of subjectivity classification is to distinguish sentences used to present opinions and other forms of subjectivity (subjective sentences set S_s) from sentences used to objectively present factual information (objective sentences set S_o), where $S_s \cup S_o = S$.

B. *Sentiment Classification*

Once the task of finding whether a piece of text is opinionated is over we have to find the polarity of the text i.e., whether it expresses a positive or negative opinion. Sentiment classification can be a binary classification (positive or negative) [8], multi-class classification (extremely negative, negative, neutral, positive or extremely positive), regression or ranking [9].

Depending upon the application of the sentiment analysis, sub -tasks of opinion holder extraction and object feature extraction are optional. (They have been represented by dashed boxes in Fig. 3).

C. *Opinion Holder Extraction*

Sentiment Analysis also involves elective tasks like opinion holder extraction, i.e. the discovery of opinion holders or sources [20, 21]. Detection of opinion holder is to recognize direct or indirect sources of opinion. They are vital in news articles and other formal documents because multiple opinions can be expressed in the same article corresponding to different opinion holders. In documents like these, the multiple opinion holders may explicitly be mentioned by name. In social networks, review sites and blogs the opinion holder is usually the author who may be identified by the login credentials.

D. *Object /Feature Extraction*

An additional task is the discovery of the target entity. In contrast with review sites, blogs and social media sites tend not have a set intention or predefined topic and are thus, inclined to discuss assorted topics. In such platforms it becomes necessary to know the target entity.

Also as mentioned before target entities can have features or components that are being reviewed. A reviewer can have differing opinions about the different features or components of the target entity. As a result,

Copyright © 2012 MECS

feature based sentiment analysis, i.e. extraction of object feature and the related opinion, is an optional task of sentiment analysis [22,23, 24].

V. Levels of Sentiment Analysis

The tasks described in the previous section can be done at several levels of granularity, namely, word level, phrase or sentence level, document level and feature level. The following Fig. 4 depicts the levels of granularity of sentiment analysis.

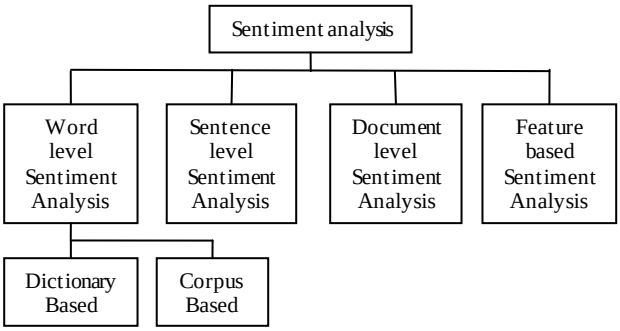

Fig.4 Granularity Levels of Sentiment analysis

The sentiment analysis tasks can be accomplished at the following levels of granularity:

A. *Document Level Sentiment Analysis*

Document-level sentiment analysis considers the whole document as the basic unit whose sentiment orientation is to be determined. To simplify the task, it is presumed that each text's overall opinion is completely held by a single opinion holder and is about a single object. Various machine learning approaches exist for this task. Pang et al. [8] used traditional machine learning methods to classify reviews as positive and negative. They experimented with three classifiers (Naive Bayes, maximum entropy, and support vector machines) and features like unigrams, bigrams, term frequency, term presence and position, and parts-of-speech. They concluded that SVM classifier works best and that unigram presence information was most effective. Document level sentiment analysis has also been formulated as a regression problem by Pang and Lee [9]. Supervised learning was used to predict rating scores. A simple and straightforward method is to find a linear combination of the polarities in the document, as given by Dave et al. [5] and Turney[25].

The difficulty lies in the fact that there could be mixed opinions in a document, and with the creative nature of natural language, people may express the same opinion in vast ways, sometimes without using any opinion words. Also as stated earlier, a text is equally likely to contain objective sentences along with subjective sentences. Therefore, tools are required to extract useful information from subjective sentences instead of objective ones. This leads to sentence level sentiment analysis.

B. *Sentence Level Sentiment Analysis*

At sentence level, research has been done on detection of subjective sentences in a document from a mixture of objective and subjective sentences and then, the sentiment orientation of these subjective sentences is determined. Yu and Hazivassiloglou [26] try to classify subjective sentences and also determine their opinion orientations. For subjective or opinion sentence identification, it uses supervised learning. For sentiment classification of each identified subjective sentence, it used a similar method to Turney[25], but with many more seed words, and log-likelihood ratio as the score function. A simple method used by Liu et al. [27], was to aggregate the orientations of the words in the sentence to get over all polarity of the opinion sentence.

One would expect that subjective sentence detection could be done by using a good sentiment lexicon, but the tricky part is that objective sentences can also contain opinion words

C. *Word Level Sentiment Analysis*

The work to find semantic orientation at phrase level is an important task of sentiment analysis. Most works use the prior polarity [28] of words and phrases for sentiment classification at sentence and document levels. Thus, the manual or semi-automatic construction of semantic orientation word lexicon is popular. Word sentiment classification use mostly adjectives as features but adverbs, and some verbs and nouns are also used by researchers [29, 30]. The two methods of automatically annotating sentiment at the word level are: (1) dictionary-based approaches and (2) corpus-based approaches.

1) *Dictionary based Methods*

In this method, a small seed list of words with known prior polarity is created. This seed list is then extended by extracting synonyms or antonyms iteratively from online dictionary sources like WordNet [1]. Kim and Hovy[31] manually created two seed lists consisting of positive and negative verbs and adjectives. They then expanded these lists by extracting, from WordNet, the synonyms and antonyms of the words of the seed list and assigning them to appropriate list (synonyms were placed in the same list and antonyms in the opposite). The sentiment strength of the words was determined by how the new unseen words interacted with the seed list. Both positive and negative sentiment strengths was computed for each word and their relative magnitudes was compared. Based on WordNet lexical relation, Kamps et al. [32] measured the semantic orientation of words.. They collected words and all their synonyms in WordNet, i.e. words of the same synset. Then a graph was created with edges connecting pairs of synonymous

[1] http://wordnet.princeton.edu/

words. The semantic orientation of a word was calculated by its relative distance from the two seed terms good and bad. The distance was the length of a shortest path between two words w_i and w_j. The values ranged from [-1, 1] with the absolute value indicating the strength of the orientation

The drawback of using a dictionary method is that the polarity classification is not domain specific. For example, "unpredictable" is a positive description for a movie plot but a negative description for a car's steering abilities [25].

2) *Corpus based Methods*

Corpus based methods rely on syntactic or statistical techniques like co-occurrence of word with another word whose polarity is known. Hatzivassiloglou and McKeown[33] predicted the orientation of adjectives by assuming that pairs of conjoined adjectives have same orientation (if conjoined by and) and opposite orientation (if conjoined by but). Thus they used conjunctions such as "corrupt and brutal" or "simplistic but well-received" to form clusters of similarly and oppositely-oriented words using a log linear regression model. They intuitively assigned the cluster that contained terms of higher average frequency as the positive list. As this method is an unsupervised classification method, the corpus required was immense. Turney [25] assigned semantic orientation by using association. That is it is said to have a positive orientation if they have good associations (e.g. Romantic ambience). The association relationship between an unknown word and a set of manually-selected seeds (like excellent and poor) was used to classify it as positive or negative The degree of association between the unknown word and the seed words was determined by counting the number of results returned by web searches in the AltaVista Search Engine joining the words with the NEAR operator and calculating the point-wise mutual information between them.

With document, sentence and phrase level analysis, we do not know what the opinion holder is expressing opinion on. Furthermore, we do not know the features that are being talked about.

D. *Feature Based Sentiment Analysis*

In a review, its author talks about the positives and negatives of a product. The reviewer may like some features and dislike some, even though the general opinion of the product may be positive or negative. This kind of information is not provided by document level or sentence level sentiment classification. Thus, feature based opinion sentiment analysis [22, 23, 24] is required. This involves extracting product feature and the corresponding opinion about it. Instinctively, one might think that product features are expressed by nouns and noun phrases, but not all nouns and noun phrases are product features. Yi et al.[29] restricted the candidate words further by extracting only base noun phrases, definite base noun phrase(noun phrases preceded by a definite article "the") and beginning definite base noun phrases(definite base noun phrase at the beginning of a sentence followed by a verb phrase). For each sentiment phrase detected, its target and final polarity is determined based on a sentiment pattern database.

Hu and Lui[30] extract the feature that people are most interested in and thus extract the most frequent noun or noun phrase using association mining. They use simple heuristic method of assigning the nearest opinion word to a feature to determine the sentiment orientation. Popescu and Etzioni[24] greatly improved the task of extracting features. They distinguish between being a part of an object and a property of the object by using WordNet's "is-a" hierarchy and morphological clues. Their algorithm tries to eliminate those noun phrases that probably are not product features. They associated meronymy discriminators with each product class and evaluated noun phrases by computing the PMI (Point-wise Mutual Information) between the phrase and meronymy discriminators.

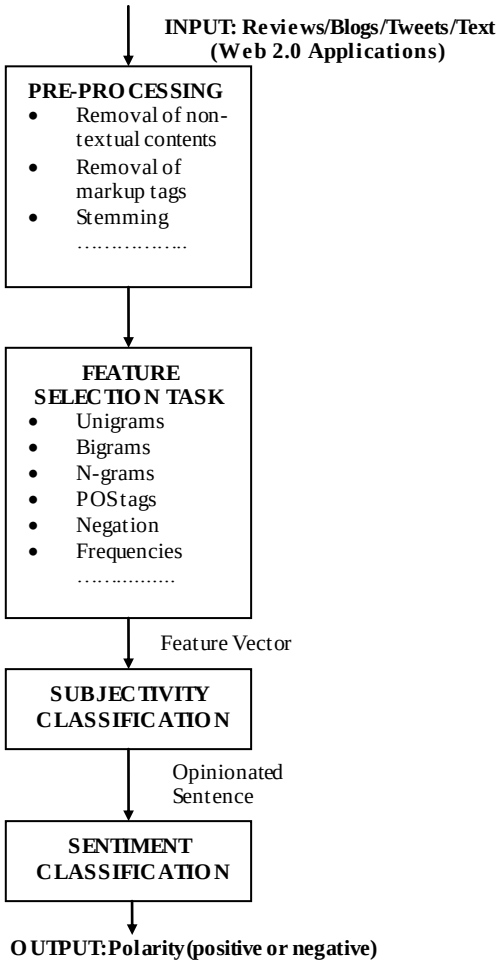

Fig.5 The Sentiment Analysis Model

Table 1 Summary of Sentiment Analysis Tasks

Sentiment Analysis Tasks
At Document Level • Task: Sentiment Classification of whole document • Classes: Positive, negative and neutral • Assumption : Each Document focuses on a single object (not true in discussion posts, blogs ,etc.) and contain opinion from a single opinion holder

At Sentence Level
• Task 1: Identifying Subjective/ Objective Sentences
  o Classes: Objective and Subjective
• Task 2: Sentiment Classification of Sentences
  o Classes: positive and negative
  o Assumption: A sentence contains only one opinion which may not always be true
 	
Prior polarities of words determined at word level sentiment analysis is used here

At Feature Level • Task 1: Identify and extract object features that have been commented on by an opinion holder (eg. A reviewer) • Task 2: Determining whether the opinions on features are negative, positive or neutral • Task 3: Find feature synonyms

VI. State-of-Art: The Past and Present of Sentiment Analysis

Most researchers have defined the Sentiment Analysis problem as essentially a text classification problem and machine learning techniques have proved their dexterity in resolving the sentiment analysis tasks [34]. Machine learning techniques require representing the key features of text or documents for processing. These key features are represented as feature vectors which are used for the classification task.. Some examples features that have been reported in literature are:

- **Words and their frequencies**

Unigrams, bigrams and n-grams along with their frequency counts are considered as features. There has been contention on using word presence rather than frequencies to better describe this feature. Pang et al. [8] showed better results by using presence instead of frequencies.

- **Parts of Speech Tags**

Parts of speech like adjectives, adverbs and some groups of verbs and nouns are good indicators of subjectivity and sentiment.

- **Syntax**

Syntactic patterns like collocations are used as features to learn subjectivity patterns by researchers. The syntactic dependency patterns can be generated by parsing or dependency trees.

- **Opinion Words and Phrases**

Apart from specific words, some phrases and idioms which convey sentiments can be used as features, e.g. “cost someone an arm and leg” [3].

- **Position of Terms**

The position of a term within a text can effect on how much the term affects overall sentiment of the text.

- **Negation**

Negation is an important but tricky feature to incorporate. The presence of a negation usually changes the polarity of the opinion but all appearances do it. For e.g., “no doubt it is the best in the market”

As we reviewed the literature for this survey, it was identified that different approaches have been applied to predict the sentiments of words, expressions or documents as to automate the sentiment analysis task. These were either a Natural Language Processing (NLP) research endeavor or addressed by Machine Learning algorithms. Our earlier work [34] probes the role of machine learning as a prominent assisting technology that has ascertained substantial gains in automated sentiment analysis research and practice by developing standards and improving effectiveness. It expounds the unique aspects of the machine learning techniques in sentiment analysis mainly because of the different features involved in case of supervised and semi-supervised techniques. Unsupervised techniques use sentiment driven pattern to obtain labels for words and phrases. While machine learning methods have established to generate good results, there are associated disadvantages. Machine learning classification relies on the training set used, the available literature reports detail classifiers with high accuracy, but they are often tested on only one kind of sentiment source, mostly movie review, thus limiting the performance indication in more general cases. Further, gathering the training set is also arduous; the noisy character of input texts and cross-domain classification add to the complexities and thus push the need for continued development in the area of sentiment analysis.

The research has further substantiated that the existing approaches to sentiment analysis can be grouped into four main categories, namely: *keyword spotting*, where the text is classified in accordance to the presence of reasonably unambiguous affect words; *lexical affinity*, defined as a probabilistic affinity for a particular emotion or opinion polarity to arbitrary words is calculated; *statistical methods*, where the significance of keywords and word co-occurrence frequencies using a large training corpus are computed ; and the most recent *sentic computing* [35], based upon a biologically-inspired and psychologically-motivated affective categorization model which makes use of ontologies and common sense reasoning tools for a conceptual-level analysis of natural language text.

Copyright © 2012 MECS

The following Table 2 depicting some previous prominent attempts to study sentiment analysis.

Table 2 Summary of the State-of-Art of Sentiment Analysis

Author	Granularity Level	Model	Features	Data Source
Hatzivassiloglou and McKeown (1997) [33]	Document	Log Linear Regression	Conjunctions and Adjectives	World Street Journal
Das and Chen (2001) [37]	Document	Lexicon and grammar rules	Words	Financial News
Pang et al. (2002) [8]	Document	NB[1], SVM[2], ME[3]	Unigram, bigram, contextual effect of negation, frequency, position	IMBD (Movie Review)
Turney (2002) [25]	Document	PMI-IR[4]	Bigrams	Automobile, bank, movie, travel reviews
Morinaga et al. (2002) [38]	Document	Decision tree induction	Characteristic words, co-occurrence words, and phrases	Cellular phones, PDA and internet service providers
Yi et al. (2003) [22]	Topic	NLP-pattern based	Feature lexical semantics	Digital camera and music reviews
Turney and Littman (2003) [39]	Document	SO-LSA[5], SO-PMI[6], General inquirer	Words and phrases	TASA-ALL corpus(from sources like novel and news articles)
Dave et al. (2003) [5]	Document	Scoring, Smoothing, NB, SVM, ME	Unigrams, bigrams and trigrams	Product reviews
Pang and Lee (2004) [40]	Document	NB, SVM	Unigram; Sentence level subjectivity summarization based on minimum cuts.	Movie Reviews
Kim and Hovy (2004) [31]	Phrase	Probabilistic based		DUC corpus
Gamon (2004) [41]	Document	SVM		Customer feedback
Nigam and Hurst (2004) [42]	Sentence	syntactic rules based chunking	Lexicon of polar phrase and their parts of speech, syntactic pattern	Usenet message board and other online resources
Pang and Lee (2005) [9]	Document	SVM, regression, Metric Labelling		Movie Reviews
Choi et al.(2005) [20]	Extract opinion holder, emotion and sentiment	CRF[7] and AutoSlog	Automatically learned extraction patterns	MQPA corpus
Wilson et al. (2005) [28]	Phrase	BoosTexter	Subjectivity Lexicon	MQPA corpus
Hu and Liu (2005) [23]	Product Feature	Opinion word extraction and aggregation enhanced with WordNet	Opinion words opinion sentences	Amazon Cnn.net
Airoldi et al. (2005) [43]	Document	Two stage Markov Blanket Classifier	Dependence among words, minimal vocabulary	IMBd, Infonic
Aue and Gamon (2005) [30]	Sentence	NB	Stemmed terms, their frequency and weights	Car reviews
Popescu and Etzioni (2005) [24]	Phrase	Relaxation Labelling Clustering	Syntactic dependency template, conjunctions and disjunctions Wordnet	Amazon Cnn.net
Cesarano, (2006) [44]	Sentence	Template based using a hybrid evaluation method	POS, n-grams	Newsarticles, web blogs
Kˆonig and Brill (2006) [45]	Document	Pattern based, SVM, Hybrid		Movie reviews, customer feedback
Kennedy and Inkpen (2006) [46]	Document	SVM, term-counting method, a combination of the two	Term frequencies	General Inquirer dictionary, CTRW dictionary & IMBd
Thomas et al. (2006) [47]	Sentence	SVM	Reference Classification	2005 U.S. floor debate in the House of Representatives

[1] Naïve Bayes
[2] Support Vector Machines
[3] Maximum Entropy
[4] Pointwise Mutual Information and Information Retrieval
[5] Semantic Orientation Latent Semantic Analysis
[6] Semantic Orientation Pointwise Mutual Information
[7] Conditional Random Field

Kaji and Kitsuregawa (2007) [36]	Phrase	Phrase trees and word co-occurrence, PMI	lexical relationships, word	HTML documents
Blitzer et al. (2007) [48]	Document	Structural Correspondence Learning	Word frequency and co-occurrence, part of speech	Book, DVD and kitchen appliance product review
Godbole et al. (2007) [49]	Word	Lexical (WordNet)	graph distance measurements between words based on relationships of synonymity and anonymity, commonality of words	Newspaper, blogpost
Annett and Kondrak (2009) [50]	Document	lexical (WordNet) & SVM	number of positive/negative adjectives/adverbs, presence, absence or frequency of words, minimum distance from pivot words	Movie review, blog posts
Zhou and Chaovalit (2008) [51]	Document	ontology-supported polarity mining	n-grams, words, word senses	Movie reviews
Hou and Li (2008) [52]	Sentence	CRF	POS tags, comparative sentence elements	Product reviews, forum discussions; labelled manually and automatically
Ferguson et al. (2009) [53]	Phrase	MNB[1]	binary word feature vectors	Financial blog articles
Tan et al. (2009) [54]	Document	NB Classifier with feature adaptation using Frequently Co-occurring Entropy	words	Education reviews, stock reviews and computer reviews
Wilson et al. (2009) [55]	Phrase	boosting, memory-based learning, rule learning, and support vector learning	words, negation, polarity modification features	MPQA Corpus
Melville et al. (2009) [13]	Document	Bayesian classification with lexicons and training documents	Words	Blog Posts, reviewing software, political blogs, movie reviews
Pak and Paroubek (2010) [15]	Sentence	MNB classifier	N-gram and POS-tags as features	Twitter posts
Barbosa and Feng (2010) [16]	Sentence	SVM	retweet, hashtags, link, punctuation and exclamation marks in conjunction with features like prior polarity of words and POS of words	Twitter posts
Heerschop (2011) [56]	document	Creates a list of adjectives and scored	POS, n-grams, negation	Text documents

[1] Multinomial Naïve Bayes

VII. Applications of Sentiment Analysis

The boom in the availability of opinionated and emotionally charged data from various review sites, blog, forums and social networks has created a wave of interest in sentiment analysis by both academia and businesses. This is because there are many practical and potential applications of sentiment analysis. Sentiment analysis assists organizations and service providers to know the mindset of their customers and users and to accordingly tailor their products and services to the needs of customers and users. It is also of vital interest for scientists such as social psychologists as it allows them to tap into the psychological thinking and responses of online communities. Following is a brief discussion on the potential applications of sentiment analysis:

A. *Bussiness Applications*

Sentiment analysis is being adopted by many businesses who would like an edge and an insight into the "market sentiment" [36]. Potential applications would be extracting product review, brand tracking, modifying marketing strategies and mining financial news. The activities that are aided by sentiment analysis are:

- Automatic tracking of combined user opinions and ratings of brands, products and services from review sites [55].
- Analyzing purchaser inclinations, competitors, and market trends
- Gauging reaction to company-related events and incidents, like during a new product launch it can give them instant feedback about the reception of the new product. It can gauge what their brand image is, whether they are liked or not.
- Monitoring crucial issues to avert harmful viral effects, like dealing with customer complaints that occur in social media and routing the complaints to the particular department that can handle it, before the complaints spread.

Key challenges identified by researchers for this application include, identifying aspects of product, associating opinions with aspects of product, identifying fake reviews and processing reviews with no canonical forms.

B. *Politics*

Sentiment analysis enables tracking of opinion on issues and subjectivity of bloggers in political blogs. Sentiment analysis can help political organization to understand which issues are close to the voter's heart [17]. Thomas et al. [47], try to determine from the transcripts of US Congressional floor debates which speeches support and which are in opposition to proposed legislation. To improve the worth of the information available to voters, the position of public figures, i.e. causes they support or oppose, can also be determined. Mullen and Malouf [58] describe a statistical sentiment analysis method on political discussion group postings to judge whether there is opposing political viewpoint to the original post. Twitter posts have been used to predict election results [59]. Researchers have collectively pointed out some research challenges namely identifying of opinion holder, associated opinion with issues, identifying public figures and legislation.

C. *Recommender System*

Recommender systems can benefit by extracting user rating from text. Sentiment analysis can be used as a sub-component technology for recommender systems by not recommending objects that receive negative feedback [60]. Pang et al. [8] classified movie reviews as "recommended" and "not recommended".

D. *Expert Finding*

There is potential of using sentiment analysis in expert finding systems. Taboada et al. [61], use sentiment analysis techniques to track literary reputation. Piao et al. [62] resolve if an author is referencing a piece of work for substantiation or as research that he or she disregards. Kumar & Ahmad [63] propose mining the expertise from the virtual community using sentiment analysis of each group member's blog & comments received on it. Their combined orientation strength determines the blog score which enables ranking the blogs and identify the expert as the one with the highest blog rank.

E. *Summarization*

Opinion summarization finds application when the number of online review of a product is large. This may make it hard for both the customer and the product manufactured. The consumer may not be able to read all the reviews and make an informed decision and the manufacturer may not be able to keep track of consumer opinion. Liu et al. [27] thus took a set of reviews on a certain product and (i) identified product features commented on (ii) identified review sentences that give opinions for each feature; and (iii) produced a summary using the discovered information. Summarization of single documents [40] or multiple documents (multiple viewpoints) [64] is also an application that sentiment analysis can augment.

F. *Government Intelligence*

Government intelligence is one more application for sentiment analysis. It has been proposed by monitoring sources, the increase in antagonistic or hostile communications can be tracked [65]. For efficient rule making, it can be used to assist in automatically analyzing the opinions of people about pending policies or government-regulation proposals. Other applications include tracking the citizen's opinion about a new scheme, predicting the likelihood of the success of a new legislative reform to be introduced and gauging the mood of the public towards a scandal or controversy.

VIII. Issues and Challenges of Sentiment Analysis

Tackling the fuzzy definition of sentiment and the complexity of its expression in text brings up new questions providing abundant opportunities for quantitative and qualitative work. Major challenges are:

A. *Keyword Selection*

Topic based classification usually uses a set of keywords to classify texts in different classes. In sentiment analysis we have to classify the text in to two classes (positive and negative) which are so different from each other. But coming up with a right set of keyword is not a petty task. This is because sentiment can often be expressed in a delicate manner making it tricky to be identified when a term in a sentence or document is considered in isolation. For example, "If you are reading this because it is your darling fragrance, please wear it at home exclusively, and tape the windows shut." (Review by Luca Turin and Tania Sanchez of the Givenchy perfume Amarige, in *Perfumes: The Guide*, Viking 2008.) No ostensibly negative words occur [4].

Copyright © 2012 MECS

B. *Sentiment is Domain Specific*

Sentiment is domain specific and the meaning of words changes depending on the context they are used in. The phrase "*go read the book*" would be considered favorably in a book review, but if expressed in a movie review, it suggests that the book is preferred over the movie, and thus have an opposite result [4].

C. *Multiple Opinions in a Sentence*

Single sentence can contain multiple opinions along with subjective and factual portions. It is helpful to isolate such clauses. It is also important to estimate the strength of opinions in these clauses so that we can find the overall sentiment in the sentence, e.g, "*The picture quality of this camera is amazing and so is the battery life, but the viewfinder is too small for such a great camera*", expresses both positive and negative opinions [4].

D. *Negation Handling*

Handling negation can be tricky in sentiment analysis. For example, "*I like this dress*" and "*I don't like this dress*" differ from each other by only one token but consequently are to be assigned to different and opposite classes. Negation words are called polarity reversers and papers [36, 45] have tried to model negation accurately. But there are many complex polarity reversers like "avoid" in *"[it] avoids all cliché's and predictability found in Hollywood movies*" [4] that have to be addressed.

E. *Sarcasm*

Sarcasm and irony are very quiet difficult to identify. Sarcasm is a very often used in social media.eg *"thank you Janet Jackson for yet another year of Super Bowl classic rock!"* (Twitter). This refers to the supposedly lame music performance in super bowl 2010 and attributes it to the aftermath of the scandalous performance of Janet Jackson in the previous year [66].

F. *Implicit Opinion*

Sentiment that appears in text can be characterized as: explicit where the subjective sentence directly conveys an opinion "*We had a wonderful time*", and implicit where the sentence implies an opinion *"The battery lasted for 3 hours"*. Present sentiment analysis models will not be able to detect this implicit opinion as a negative opinion.

G. *Comparative Sentences*

A comparative sentence expresses a relation based on similarities or differences of more than one object [3]. Research on classifying a comparative sentence as opinionated or not is limited. Also the order of words in comparative sentences manifests differences in the determination of the opinion orientation. E.g. The sentence, "*Car X is better than Car Y*" communicates a completely opposite opinion from "*Car Y is better than Car X*".

H. *Multilingual Sentiment analysis*

Most sentiment analysis research has focused on data in the English language, mainly because of the availability of resources like lexicons and manually labeled corpora. As only 26.8 % of Internet users speak English[1], the construction of resources and tools for subjectivity and sentiment analysis in languages other than English is a growing need. Several methods have been proposed to leverage on the resources and tools available in English by using cross-lingual projections [67].

I. *Opinion Spam*

Opinion spam refers to fake or bogus opinions that try to deliberately mislead readers or automated systems by giving undeserving positive opinions to some target objects in order to promote the objects and/or by giving malicious negative opinions to some other objects in order to damage their reputations [3]. Many review aggregation sites try to recognize opinion spam by procuring the helpfulness or utility score of each review from the reader by asking them to provide helpfulness feedbacks to each review ("Was this review helpful?").

IX. Conclusion

This paper illustrates the research area of Sentiment Analysis and its latest advances. It affirms the terminology, the major tasks, the granularity levels, and applications of sentiment analysis. It also discusses the impact of Web 2.0 applications on this research field. Most work has been done on product reviews – documents that have a definite topic. More general writing with varied domains, such as blog posts, tweets, posts and web pages, have recently been creating & receiving attention. Future work in expanding existing techniques to handle more general writings and crossing domains is an exciting opportunity for both academia and businesses.

References

[1] Tim O'Reilly, Web 2.0 Compact Definition: Trying Again (O'Reilly Media, Sebastopol), http://radar.oreilly.com/archives/2006/12/web_20_compact.html. Accessed 22 Mar 2007

[2] Tang H, Tan S, and Cheng X. A survey on sentiment detection of reviews. Expert Systems with Applications: An International Journal, September 2009, 36(7):10760–10773.

[3] Liu B. Sentiment Analysis and Subjectivity. Handbook of Natural Language Processing, Second edition, 2010

[4] Pang, B and Lee L. Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieva,l 2008,(1-2),1–135

[1] http://www.internetworldstats.com/stats7.htm

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

[6] Wiebe, J., Wilson, T., Bruce, R., Bell, M., and Martin, M. Learning subjective language. Computational Linguistics, 2004, 30(3):277–308

[7] Theresa Wilson, Janyce Wiebe, and Rebecca Hwa. Just how mad are you? Finding strong and weak opinion clauses. In Proceedings of AAAI, 2004,pages 761–769.

[8] Pang, B., Lee, L., and Vaithyanathan.S. Thumbs up? Sentiment classification using machine learning techniques. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2002, (EMNLP):79–86.

[9] Pang B. and Lee L. Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales, Proceedings of the Association for Computational Linguistics (ACL),2005:115–124

[10] Anderson, P. What is Web 2.0? Ideas, technologies and implications for education. Technical report, JISC,2007

[11] Mishne G. and Glance N. Predicting movie sales from blogger sentiment. In AAAI Symposium on Computational Approaches to Analyzing Weblogs (AAAI-CAAW),2006: 155–158.

[12] Liu, Y., Huang, J., An, A., and Yu, X. ARSA: A sentiment-aware model for predicting sales performance using blogs. In Proceedings of the ACM Special Interest Group on Information Retrieval (SIGIR),2007

[13] Melville, P., Gryc, W., and Lawrence, R.D. Sentiment analysis of blogs by combining lexical knowledge with text classification. In Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining.2009: 1275-1284.

[14] Go, A., Bhayani, R., Huang, L. Twitter sentiment classification using distant supervision. Technical report, Stanford Digital Library Technologies Project.2009.

[15] Pak A. and Paroubek P. Twitter as a corpus for sentiment analysis and opinion mining. In Proceedings of the Seventh Conference on International Language Resources and Evaluation, .2010:1320-1326.

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

[17] O'Connor B., Balasubramanyan R., Routledge B.R., Smith N. A. From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. AAAI. 2010

[18] Hatzivassiloglou, V. and Wiebe, J. Effects of adjective orientation and gradability on sentence subjectivity. In Proceedings of the International Conference on Computational Linguistics (COLING), 2000.

[19] Riloff E. and Wiebe J., Learning extraction patterns for subjective expressions.Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP) , 2003.

[20] Choi, Y., Cardie, C., Riloff, E., and Patwardhan, S., Identifying sources of opinions with conditional random fields and extraction patterns. Proceedings of the Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP), 2005.

[21] Bethard, S., Yu, H., Thornton, A., Hatzivassiloglou, V., and Jurafsky, D., Automatic extraction of opinion propositions and their holders. Proceedings of the AAAI Spring Symposium on Exploring Attitude and Affect in Text, 2004.

[22] Yi, J., Nasukawa, T., Niblack, W., &Bunescu, R.,Sentiment analyzer: extracting sentiments about a given topic using natural language processing techniques. Proceedings of the 3rd IEEE international conference on data mining (ICDM 2003):427–434

[23] Hu, M. and Liu, B. Mining opinion features in customer reviews. In Proceedings of AAAI, 2004: 755–760.

[24] Popescu A-M. and Etzioni O., Extracting product features and opinions from reviews, Proceedings of the Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP),2005

[25] Turney P. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In Proceedings of the Association for Computational Linguistics (ACL), 2005: 417–424.

[26] Yu H. and Hatzivassiloglou V., Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences.In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2003.

[27] Liu, B., Hu, M., & Cheng, J. Opinion observer: Analyzing and comparing opinions on the web. In Proceedings of the 14th international world wide web conference (WWW-2005). ACM Press: 10–14.

[28] Wilson T., Wiebe J., and Hoffmann P.,Recognizing contextual polarity in phrase-level sentiment analysis. In Proceedings of the Human Language Technology Conference and the Conference on Empirical Methods in Natural

Language Processing (HLT/EMNLP)2005: 347–354

[29] Esuli, A., & Sebastiani, F. Determining the semantic orientation of terms through gloss classification.In Proceedings of CIKM-05, the ACM SIGIR conference on information and knowledge management, Bremen, DE, 2005.

[30] Aue, A. and Gamon, M., Customizing sentiment classifiers to new domains: A case study. Proceedings of Recent Advances in Natural Language Processing (RANLP),2005.

[31] Kim, S. and Hovy, E., Determining the sentiment of opinions.In Proceedings of the International Conference on Computational Linguistics (COLING) ,2004

[32] Kamps, J., Marx, M., Mokken, R.J., de Rijke, M., Using WordNet to measure semantic orientation of adjectives. In Language Resources and Evaluation (LREC),2004.

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

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

[35] Cambria, Erik. Roelandse, Martijn. ed. Sentic Computing: Techniques, Tools and Applications. Berlin: Springer-Verlag., 2012.

[36] Kaji, N. and Kitsuregawa, M., Building lexicon for sentiment analysis from massive collection of html documents. Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2007.

[37] Das, S. and Chen, M., Yahoo! for Amazon: Extracting market sentiment from stock message boards. In Proceedings of the Asia Pacific Finance Association Annual Conference (APFA),2 001.

[38] Morinaga S., Yamanishi K., Tateishi K., and Fukushima T., Mining product reputations on the web.In Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2002: 341–349, Industry track

[39] Turney, P. and Littman, M., Measuring praise and criticism: Inference of semantic orientation from association. ACM Transactions on Information Systems (TOIS),2003, 21(4):315–346.

[40] Pang B. and Lee L., A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In Proceedings of the Association for Computational Linguistics (ACL), 2004: 271–278.

[41] Gamon, M,. Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis. In Proceedings of the International Conference on Computational Linguistics (COLING), 2004.

[42] Nigam, K. and Hurst, M.,.Towards a robust metric of opinion.The AAAI Spring Symposium on Exploring Attitude and Affect in Text,2004

[43] Airoldi, E. M., Bai, X., and Padman, R., Markov blankets and meta-heuristic search: sentiment extraction from unstructured text. Lecture Notes in Computer Science,2006, 3932: 167–187.

[44] Cesarano, C., Dorr, B., Picariello, A., Reforgiato, D., Sagoff, A., Subrahmanian, V.: OASYS: An Opinion Analysis System. AAAI Press In: AAAI Spring Symposium on Computational Approaches to Analyzing Weblogs (CAAW 2006): 21–26.

[45] K¨onig, A. C. & Brill, E., Reducing the human overhead in text categorization. In proceedings of the 12th ACM SIGKDD conference on knowledge discovery and data mining,2006, pp: 598–603.

[46] Kennedy, A. and Inkpen, D., Sentiment classification of movie reviews using contextual valence shifters. Computational Intelligence, 22(2, Special Issue on Sentiment Analysis), 2006: 110–125.

[47] Thomas M, Pang B., and Lee L., Get out the vote: Determining support or opposition from Congressional floor-debate transcripts. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2006: 327–335.

[48] Blitzer, J., McDonald, R., and Pereira, F., Domain adaptation with structural correspondence learning. In Empirical Methods in Natural Language Processing (EMNLP), 2006.

[49] Godbole, N., Srinivasaiah, M., and Skiena, S., Large-scale sentiment analysis for news and blogs. Proceedings of the International Conference in Weblogs and Social Media, 2007.

[50] Annett, M. and Kondrak, G. A comparison of sentiment analysis techniques: Polarizing movie blogs. Advances in Artificial Intelligence,2008, 5032:25–35.

[51] Zhou, L. and Chaovalit, P., Ontology-supported polarity mining.Journal of the American Society for Information Science and Technology,2008, 69:98–110.

[52] Hou, F. and Li, G.-H., Mining chinese comparative sentences by semantic role labeling. Proceedings of the Seventh International Conference on Machine Learning and Cybernetics, 2008.

[53] Ferguson, P., O'Hare, N., Davy, M., Bermingham, A., Tattersall, S., Sheridan, P., Gurrin, C., and Smeaton, A. F., Exploring the use of paragraph-level annotations for sentiment analysis in financial blogs.1st Workshop on Opinion Mining and Sentiment Analysis (WOMSA),2009.

[54] Tan, S., Cheng, Z., Wang, Y., and Xu, H., Adapting naive bayes to domain adaptation for sentiment analysis. Advances in Information Retrieval, 2009, 5478:337–349.

[55] Wilson, T., Wiebe, J., and Hoffmann, P., Recognizing contextual polarity: an exploration of features for phrase-level sentiment analysis. Computational Linguistics, 2009, 35(5):399–433

[56] Heerschop, B., Hogenboom, A., and Frasincar, F. Sentiment lexicon creation from lexical resources, Springer, In 14th International Conference on Business Information Systems (BIS 2011), volume 87 of Lecture Notes in Business Information Processing: 185–196.

[57] Chen Y. Y. and Lee K. V., User-Centered Sentiment Analysis on Customer Product Review. World Applied Sciences Journal 12 (Special Issue on Computer Applications & Knowledge Management),2011: 32-38

[58] Mullen T. and Malouf R.,Taking sides: User classification for informal online political discourse. Internet Research, 2008, 18:177–190.

[59] Tumasjan A., Sprenger T.O., Sandner P.G., Welpe I. M., Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. AAAI ,2010

[60] Terveen L., Hill W., Amento B., McDonald D., and Creter J., PHOAKS: A system for sharing recommendations. In Communications of the Association for Computing Machinery (CACM),2007, 40(3):59–62

[61] Taboada M., Gillies M. A., and McFetridge P., Sentiment classification techniques for tracking literary reputation.In LREC Workshop: Towards Computational Models of Literary Analysis, 2006: 36–43.

[62] Piao S., Ananiadou S., Tsuruoka Y., Sasaki Y., and McNaught J., .Mining opinion polarity relations of citations. In International Workshop on Computational Semantics 84 (IWCS), 2007:366–371.

[63] Kumar, A. & Ahmad, N. ComEx Miner: Expert Mining in Virtual Communities, International Journal of Advanced Computer Science and Applications (IJACSA), Vol.3, No. 6, June 2012, The Science and Information Organization Inc, USA.

[64] Seki Y., Eguchi K., Kando N., and Aono M., Multi-document summarization with subjectivity analysis at DUC 2005. In Proceedings of the Document Understanding Conference (DUC).

[65] Spertus E., Smokey: Automatic recognition of hostile message. In Proceedings of Innovative Applications of Artificial Intelligence (IAAI),1997: 1058–1065.

[66] Davidov, D., Tsur, O., and Rappoport, A., Semi-supervised recognition of sarcastic sentences in twitter and amazon. In Conference on Natural Language Learning (CoNLL), 2010.

[67] Denecke, K.., Using SentiWordNet for Multilingual Sentiment Analysis .Proc. of the IEEE 24th International Conference on Data Engineering Workshop (ICDEW 2008), IEEE Press:507-512.

Authors' Profile

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 Engineering; she is currently working as a scholar in the field of Sentiment Analysis.

How to cite this paper: Akshi Kumar, Teeja Mary Sebastian,"Sentiment Analysis: A Perspective on its Past, Present and Future", IJISA, vol.4, no.10, pp.1-14, 2012. DOI: 10.5815/ijisa.2012.10.01