A research study of sentiment analysis and various techniques of sentiment classification

Gaurav Dubey* and Ajay Rana

Amity University,
Sec 125, Noida UP, India
Fax: +91(0120)-4392315
E-mail: gdubey@amity.edu
E-mail: gdubey1977@gmail.com
E-mail: ajay_rana@amity.edu
*Corresponding author

Jayanthi Ranjan

Institute of Management Technology,
Raj Nagar, Post Box No. 137,
Ghaziabad U.P., 201001, India
Fax: +91-(0120)-2827895
E-mail: jranjan@imt.edu

Abstract: Sentiment analysis or opinion mining is a deterministic technique for classifying and evaluating other people's opinions. Nowadays, people build their perception and make decisions by analysing the facts and reviews of other people either manually or computationally. Since everything is online nowadays, internet has become an integrated part of human lives and is thus used for exchanging all aspects of human life viz. sentiments, emotions, affection, support, opinions, trade, business, etc. With the onset of social media there has been numerous platform such as *blogs, discussion forums, reviews and social networks* where an individual can post his or her reviews, feedbacks and list their likes and dislikes for a product's attributes or features or comparison of different products (same or different feature). These reviews are gathered and are analysed to evaluate the overall orientation of the collected reviews. This survey paper focuses on in-depth study of the topic and discusses all concepts and terminologies of opinion mining. This paper also discusses the methods and techniques used for gathering reviews, extracting the phrases based on the subjectivity (Esuli and Sebastiani, 2006) and thereafter calculating the semantic orientation of the collected reviews.

Keywords: opinion mining; sentiment analysis; sentiment classification; word of mouth; WOM; natural language processing; semantic orientation.

Reference to this paper should be made as follows: Dubey, G., Rana, A. and Ranjan, J. (2016) 'A research study of sentiment analysis and various techniques of sentiment classification', *Int. J. Data Analysis Techniques and Strategies*, Vol. 8, No. 2, pp.122–142.

Biographical notes: Gaurav Dubey received his MTech (Computer Science and Engineering) from IETE, New Delhi, BTech in Computers from Aligarh Muslim University (1999), India. He is pursuing his PhD in Computer Science in the area of data mining. He is currently employed as an Assistant Professor with Amity University Uttar Pradesh, Noida, India and has 14 years of experience in IT field. His research interest includes data mining and its applications.

Ajay Rana is the Group Director in Amity School of Engineering. He received his PhD and MTech in Computer Science and Engineering. He has published more than 135 research papers in reputed journals and proceedings of international and national conferences. He has co-authored four books and co-edited 15 conference proceedings. He has delivered invited lectures in more than 30 technical and management workshop/conferences programmes in India and abroad. He has organised 45 plus conference, workshop, faculty development programme and attend more than 27 advance courses for academics and executive development programme with respect to latest technologies.

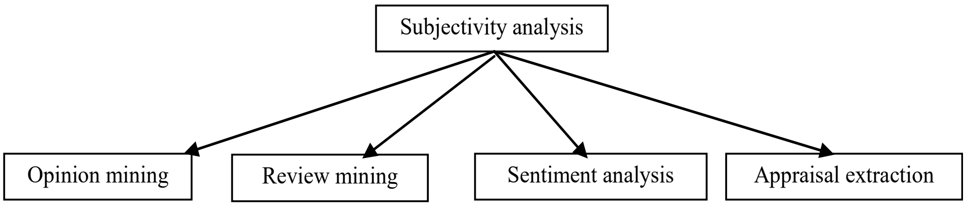
Jayanthi Ranjan is a Professor of IT and the Associate Dean of International Relations, IMT Ghaziabad and has more than 22 years of teaching, consulting and research experience in big data, business analytics and business intelligence. She is a Gold Medallist and university rank holder in her graduation. She has over 190 publications that have appeared in international refereed journals and conference proceedings and written six cases that have appeared in case houses like case centre. She has received two best paper awards. She has received many awards like Best Teacher Award from AIMA-Ghaziabad Management Association, India, Best Woman Management Researcher Award in 2009, Best Management Researcher Award in 2010, Best Management Scholar Award in 2011, Best Management Researcher Award in 2012, Global Management Research Award in 2013 from AIMS International. She is also serving at the capacity of editor in several refereed journals. She has also published ten edited books.

1 Introduction

Sentiment analysis is the part of subjectivity analysis (Akkaya et al., 2009) which is also very popular by the name opinion mining. Opinion mining is mainly concern with analysis of linguistic natural expression of individual's opinion about certain product or any other area where public opinion or review matter a most. Subjectivity analysis aims at determining the attitude of the writer or author of opinion with respect to some topic or product or services or the overall contextual polarity or tonality of a document or review (Hassan and Radev, 2010). The attitude may involve the user's experience, evaluation, judgment, the emotional state or intended emotional effect. It is a natural language processing (Indurkhya and Damerau, 2010) and information extraction task that identifies the writer's feelings and experiences expressed in positive and negative comments, questions and requests, by analysing monstrous amount of information available over the web. The major force behind the emergence of opinion mining today, is the exponential increase in internet usage and exchange or share of public views and opinions (Dellarocas et al., 2007).

The analysis of opinions may be topic-based (Jijkoun et al., 2010) where documents are classified into predefined topic classes, e.g., science, sports, entertainment, politics, etc. In topic-based classification topic-related words plays a significant role. However, in Sentiment classification (Li et al., 2010a) they are least considered. Here, the classification is at document-level, where whole document is classified based on its polarity,
i.e., sentiment words that indicate positive or negative opinions (sometimes neutral) are important, e.g., great, poor, excellent, bad, disgusting, etc. This classification can also be extended to sentence level, comparative sentence (Jindal and Liu, 2006a, 2006b), i.e., to classify each sentence as expressing a positive, negative or neutral opinion (Sugato et al., 2013).

Figure 1 Names often interchangeably used



To enable the above visualisation, identify product's review's phrases that customers have expressed their (positive or negative) opinions on. These opinions consist of user's viewpoint, fancy, attitude, sensibility, etc. The reviews can be of product's feature, its attributes or it could contain the comparison of different products of same realm (Li et al., 2010b). This is termed as opinion extraction. For each phrase, identify whether the opinion from each reviewer is positive or negative and project the overall semantic orientation, if any.

Fully analysing and classifying opinions involves tasks that relate to some fairly deep semantic and syntactic analysis of the text. These include not only recognising that the text is subjective, but also determining what the opinion is about, and which of many possible positions the holder of the opinion expresses regarding that subject. In this report, we are presenting the components of our opinion detection and organisation subsystem. These components deal with the initial tasks of classifying articles as mostly subjective or objective (Wiebe and Riloff, 2005), finding opinion sentences in both kinds of articles, and determining, in general terms and without reference to a specific subject, if the opinions are positive or negative.

First we will discuss the basic concept of subjectivity analysis and other related terms opinion mining incorporates three mining tasks. These tasks include sentiment classification (Pang et al., 2002), feature-based opinion mining and summarisation (Meng and Wang, 2009) and lastly comparative sentence and relation mining. Since web is a huge warehouse of structured and unstructured data, therefore, the most challenging aspect of opinion mining is to analyse this data and extract correct and relevant opinion and their classification.

The remainder of the paper is organised as follows. Section 2 discusses the related work that has been done so far in this area of sentiment analysis and opinion mining. Section 3 introduces the various categories of sentiment analysis definition of opinion mining and various task involved in sentiment analysis. Section 4 discuss about various machine learning techniques that are used for sentiment classification and in Section 5, we introduce aspect-based semantic classification at document level and method of opinion extraction. Sections 6.1 and 6.2 discuss feature-based opinion mining and comparative sentence mining. Finally, Section 7 concludes the paper.

2 Related work and significance

Sentiment analysis or opinion mining is a comparatively new research topic in recent years. The early work of sentiment detection began in late 1990s (Kessler et al., 1997; Spertus, 1997; Argamon-Engelson et al., 1998), but only in the early 2000s it become a major subfield of the information management discipline (Kobayashi et al., 2004). Most of the work has focused on various product reviews, there are applications to other domains such as debates (Thomas et al., 2006; Lin et al., 2006), news (Devitt and Ahmad, 2007) and blogs (Ounis et al., 2008).

Some of the other researcher who started work in sentiment analysis are Rauber and Muller-Kogler (2001), Subasic and Huettner (2001), Tong (2001), Dimitrova et al. (2002), Durbin et al. (2003), Hillard et al. (2003), Efron (2004) and Gamon et al. (2005), Glance et al. (2004) and Grefenstette et al. (2004).

Dini and Mazzini (2002) studied customer views about a product from web and applied syntactic and semantic processing to these in order to provide a structured input from natural text, for later processing with data mining algorithms. Their approach largely consisted of shallow syntactic parsing with a method known as chunking, coupled with a semantic parsing algorithm utilising a parallel template system (i.e., an advanced regular expression matcher) applying a sentiment polarity to sentences parsed.

Morinaga et al. proposed a framework to ease opinion mining, based on automatic opinion labelling in order to assess an entity's reputation. The system was built around extracting characteristic words using stochastic complexity, in order to gain a sense of overall features. A set of user specified categories are then tagged with meta-scores indicating how likely an opinion is to be expressed with a typical sentence. This is done via Bayesian theory. Finally, the found features are mapped to the opinions of the specified categories via principal component analysis.

Peter D. Turney published a paper in which they proposed an unsupervised learning algorithm for classifying reviews as recommended or not recommended. The algorithm used part-of-speech (POS) tagging to identify adjectives and adverbs, and then uses an algorithm employing point-wise mutual information along with information retrieval to measure the similarity of found opinion-bearing words with known opinion reference words (e.g., excellent for positive). For each review, the similarity scores of all the opinion-bearing words are averaged, and if it is largely positive, the review is given a positive stamp, negative if opposite. Turney (2002) achieved 74% average accuracy, on reviews ranging from automobiles to movies over banks and travel destinations.

Pang and Lee (2002) studied sentiment classification in the movie review domain. They employed a method which has no prior knowledge or training. Pang et al. used very simple system of machine learning in their paper. This method was employed by the authors for conducting a study of the performance of human-chosen words. This work was carried out further by two computer science graduate students. They critically examined existing methods of using human-annotated seed words as training input for sentiment analysis, such as that employed by Turney the same year. The two students proposed a word list of positive and negative words, respectively. They then matched these with a list composed from applying introspection and simple statistics of the data. The list proved significantly more reliable than the ones using human proposed seed words. Although this study was limited in size, indicated that automatic supervised feature selection could produce features that can be better than those of human guess-based selection. Pang et al. then utilised this knowledge in creating a very simple, entirely introspection-based opinion classifier method using machine learning.

Pang and Lee (2008) wrote a book that presents a thorough overview of the research in the field. Pang et al. (2002) conducted early polarity classification of reviews using supervised approaches. The techniques which they explored are support vector machines (SVMs), naïve Bayes (NB) and maximum entropy classifiers; they used datasets with a different set of features, such as unigrams, bigrams, binary and term frequency feature weights and others. The outcome of their observation was that sentiment classification is not that easy than standard topic-based classification they also concluded that using a SVM classifier with binary unigram-based features produces the best results. A subsequent innovation was the detection and removal of the objective parts of documents and the application of a polarity classifier on the rest (Pang and Lee, 2004). This exploited text coherence with adjacent text spans which were assumed to belong to the same subjectivity or objectivity class. Documents were represented as graphs with sentences as nodes and association scores between them as edges. Two additional nodes represented the subjective and objective poles. The weights between the nodes were calculated using three different, heuristic decaying functions. Finding a partition that minimised a cost function separated the objective from the subjective sentences. They reported a statistically significant improvement over a NB baseline using the whole text but only slight increase compared to using a SVM classifier on the entire document (Pang and Lee, 2008).

Mullen and Collier (2004) used SVMs and expanded the feature set for representing documents with favourability measures from a variety of diverse sources. They introduced features based on Osgood's et al. (1967) theory of semantic differentiation using WordNet to derive the values of potency, activity and evaluative of adjectives and Turney's (2002) semantic orientation. Their results showed that using a *hybrid* SVM classifier, which uses as features the distance of documents from the separating hyperplane, with all the above features produces the best results (Mullen and Collier, 2004).

Whitelaw et al. (2005) added fine-grained semantic distinctions in the feature set. Their approach was based on a lexicon created in a semisupervised fashion and then manually refined It consists of 1,329 adjectives and their modifiers categorised under several taxonomies of appraisal attributes based on Martin and White's (2005) appraisal theory. They combined the produced appraisal groups with unigram-based document representations as features to a SVM classifier (Witten and Frank, 1999), resulting in significant increases in accuracy.

Zaidan et al. (2007) introduced 'annotator rationales', i.e., words or phrases that explain the polarity of the document according to human annotators. By deleting rationale text spans from the original documents they created several *contrast* documents and constrained the SVM classifier to classify them less confidently than the originals. Using the largest training set size, their approach significantly increased the accuracy on a standard dataset.

Prabowo and Thelwall (2009) proposed a *hybrid* classification process by combining in sequence several ruled-based classifiers with a SVM classifier. The former were based on the General Inquirer lexicon (Wilson et al., 2005), the MontyLingua POS tagger and cooccurrence statistics of words with a set of predefined reference words. Their experiments showed that combining multiple classifiers can result in better effectiveness than any individual classifier, especially when sufficient training data is not available.

The two main popular approaches to sentiment detection, especially in the real-world applications, were based on machine learning techniques and based on semantic analysis techniques. After that, the shallow nature language processing techniques were widely used in this area, especially in the document sentiment detection.

Lexicon-based methods rely on a sentiment lexicon, a collection of known and precompiled sentiment terms. The major contribution was given by Popescu and Etzioni (2005), Scharl and Weichselbraun (2008) and Taboada et al. (2011). Machine learning approaches make use of syntactic and/or linguistic features (Pak and Paroubek, 2010). There also exists hybrid approach, with sentiment lexicons playing a key role in the majority of methods, e.g., Diakopoulos et al. (2010). For example, Moghaddam and Popowich (2010) establish the polarity of reviews by identifying the polarity of the adjectives that appear in them, with a reported accuracy of about 10% higher than pure machine learning techniques. However, such relatively successful techniques often fail when switched to new realms or context types, due to the inflexibility in the ambiguity of sentiment terms. The sentiment terms may indicate subjectivity, but there may be insufficient context to calculate its semantic orientation, particularly for adjectives in sentiment lexicons (Mullaly et al., 2010). Several evaluations have shown the significance of contextual information (Weichselbraun et al., 2010; Wilson et al., 2009), and have identified context words with a high impact on the polarity of ambiguous terms. For example, the adjective unpredictable may have a negative orientation in an automotive review, in a phrase such as ‘unpredictable steering’, but it could have a positive orientation in a movie review, in a phrase such as ‘unpredictable plot’. Therefore two consecutive words are extracted, where one member of the pair is an adjective or an adverb and the second provides context.

Recently, techniques for opinion mining have started focusing on various social medias, in combination with a trend towards its application as a proactive rather than a reactive mechanism. Understanding public opinion can have important consequences for the prediction of future events and trends.

One of the most obvious application of this is for review rating: Turney (2002) found that, with 410 reviews from opinions, the algorithm attains an average accuracy of 74%. It appears that movie reviews are difficult to classify, because the whole is not necessarily the sum of the parts; thus the accuracy on movie reviews is about 66%. On the other hand, for banks and automobiles, it seems that the whole is the sum of the parts, and the accuracy is 80% to 84%. Travel reviews are an intermediate case.

Another applications of this is for stock market predictions: Bollen and Mao (2011) found that, contrary to the expectation that if the stock markets fell, then public mood would also become more negative, in fact a drop in public mood acts as a precursor to a fall in the stock market.

Almost all the work on opinion mining from Twitter has used machine learning techniques. Pak and Paroubek (2010) aimed to classify arbitrary tweets on the basis of positive, negative and neutral sentiment, constructing a simple binary classifier which used n-gram and POS features, and trained on instances which had been annotated according to the existence of positive and negative emoticons.

Their approach has much in common with an earlier sentiment classifier constructed by Go et al. (2009), which also used unigrams, bigrams and POS tags, though the former demonstrated through analysis that the distribution of certain POS tags varies between positive and negative posts. One of the reasons for the relative paucity of linguistic techniques for opinion mining on social media is most likely due to the difficulties in using NLP on low quality text, something which machine learning techniques can – to some extent – bypass with sufficient training data. For example, the Stanford NER drops from 90.8% to 45.88% when applied to a corpus of tweets (Liu, 2010). Ritter et al. (2011) also demonstrate some of the difficulties in applying traditional POS tagging, chunking and named entity recognition techniques to tweets, proposing a solution based on labelled LDA.

Opinion mining can be useful in several ways. For example, in marketing it helps in judging the success of an ad campaign or new product launch, to determine which versions of a product or service are popular and even identify which demographics like or dislike particular features. This opinion classification is useful to both potential customers (buyers) and product manufacturers.

For a potential customer, although he/she can read all reviews of different products at merchant sites to mentally compare and assess the strengths and weaknesses of each product in order to decide which one to buy, it is much more convenient and less time consuming to see a visual feature-by-feature opinion in the reviews. A system like ours can be installed at a merchant site that has reviews so that potential buyers can compare not only prices and product specifications (which can already be done at some sites), but also opinions from existing customers. For a product manufacturer, comparing consumer opinions of its products and those of its competitors to find their strengths and weaknesses is crucial for marketing intelligence and for product benchmarking. This is typically done manually now, which is very labour intensive and time consuming. Our system comes to help naturally in this case.

3 Sentiment analysis

Sentiment analysis is recent and popular research topic in this segment we will explain some detail knowledge about sentiment analysis. Sentiment analysis is highly challenging research area of NLP, some work has been started before year 2000 but major work of sentiment analysis was started after year 2000 as we have discussed in above mentioned related work. Sentiment analysis can be categorised in main three levels, document level, sentence level, entity and aspect level.

Figure 2 Various categories of sentiment analysis

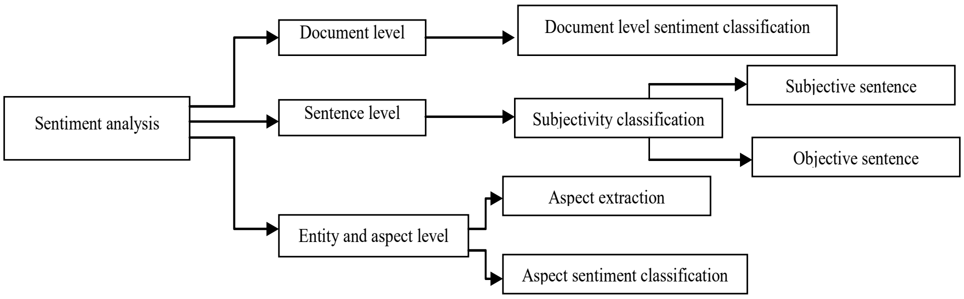


Figure 2 diagram shows a detail classification of sentiment analysis which can be explained at various levels like document level, sentence level or entity and aspect level. In the next section we will discuss basic terms related to sentiment analysis.

3.1 What is an opinion?

To understand this we will take a simple review of i-phone

> "The iPhone 5 is little expensive to other smartphone in the same segment but it is a great piece of hardware accompanied by an operating system which has started to show its age. Apple makes it easy for consumers by launching just one new phone per year while the competition launches hundreds of devices every year. If you are using an iPhone 4S, you could give the iPhone 5 a pass but if you are using an iPhone 4 or any previous generation iPhone the iPhone 5 is a must have upgrade. Then again if you are one of those who love to get your hands on the best phones first, go ahead and get the iPhone 5. Good luck finding one. They are in great demand."

There are many things that we want to mine from this review, like the various statement in the review expressing positive or negative view for the product for example "The iPhone 5 is little expensive to other smartphone" expressing some negative opinion about the phone while some sentence in the review expressing the positive view. Now in order to better understand what is an opinion we will discuss some of the definition proposed by Liu (2010). According to the Bing Liu the opinion in a sentence or a document can be defined in following below mentioned ways.

1 An opinion can be defined as consisting of two key components a target g and sentiment s, i.e., O(g, s), where g can be any entity or aspect, and s can be either a positive or negative or neutral sentiment in the document.

2 Further he defined that an opinion can also be defined as a quadruple O(g, s, h, t), where g is the opinion target, h is the opinion holder and t is the time when opinion was expressed.

3 Another definition of opinion was defined in which an opinion was presented as quintuple (e, a, s, h, t) where e is the entity, a is an aspect of e, s is the sentiment of aspect, h is the opinion holder and t is the time when opinion is expressed by h.

The last definition regarding opinion gives more precise view, the five component presented in quintuple are essential and all having their important significance in order to explain an opinion. The opinion definition provides a way to transform unstructured text data into structured data. In other words we can explain quintuple as database schema. Once the data gets the structured form the analysis on the data becomes comparatively easier.

Opinion can be categorised as regular opinion or comparative opinion. First we will discuss how sentiment analysis can be done for regular opinions.

3.2 Various task in sentiment analysis

The main objective of sentiment analysis is to first identify all opinion quintuple (e, a, s, h, t) in a document d. These five component of quintuple are used to derive key tasks of sentiment analysis. In the quintuple representation of opinion the first component is entity. The extraction of entity from the opinion is an important task, some entities in an opinion can be represented by different name, we need to recognise that they all refer to same entity. Second component of opinion is aspect of entity for example *Picture quality of phone*, in this sentence picture is the aspect of entity phone. Aspect may also be represented by different name, e.g., picture, photo, image are same aspect for phone so the next task after entity identification is aspect categorisation. Aspect expression are usually noun and noun phrases, adjective and adverbs. Aspect expression can be of two types explicit aspect expression (noun or noun phrase), e.g., in the sentence *Sound quality of this phone is excellent* sound quality is explicit aspect expression. The aspect expression which are not noun or noun phrase are called implicit aspect expression, e.g., "this mobile phone is cheap" here cheap is implicit aspect expression.

Third important component in opinion definition is sentiment, the sentiment of an aspect can be positive, negative or neutral. The other two component of opinion are opinion holder and time. Opinion holder (Liu, 2010) can be person or organisation. On the basis of above discussion a model of entity and model of opinion document is presented. Following are the steps for sentiment analysis.

Figure 3 Various steps of sentiment analysis

Step 1: Extract all entity expression in the given document D.
Step 2: Find all the aspect and their synonyms also into a cluster for that entity.
Step 3: Extract opinion holder for opinion from text or structured data.
Step 4: Extract time when the opinion was made.
Step 5: Find the sentiment polarity like positive, negative or neutral.

4 Various techniques of sentiment classification

In the previous section we explained the process of sentiment analysis, Figure 3 shows the various steps involved in the process of sentiment analysis. In this section we will discuss various sentiment classification techniques. In the earlier section we have explained that how an opinion can be represented in a document in the form of entity (e) and sentiment (s). The problem of sentiment classification can be discussed in two ways, if s takes categorical value then this task is under classification problem, if s takes numeric value then it becomes problem of regression. First we discuss sentiment classification, which implies classifying the tagged reviews into classes of polarity: positive and negative. Positive polarity implies positive orientation of review and negative polarity implies negative orientation of review. The basic approach in classifying opinion is to treat the problem as a topic-based text classification problem, then any text classification algorithm can be applied to determine the semantic orientation of the tagged reviews, such as naïve Bayesian, SVM or kNN (Yugowati et al., 2013). The orientation can also be determined using score function. We discuss three main approaches, naïve Bayesian, SVM and maximum entropy. The approach was experimented by Pang et al. using movie reviews of two classes, positive and negative. It was shown that using a unigram (a bag of individual words) in classification performed well using either naïve Bayesian or SVM. Test results using 700 positive reviews and 700 negative reviews showed that these two classification algorithms achieved 81% and 82.9% accuracy respectively with three-fold cross validation.

4.1 Classification using naïve Bayesian

An NB classifier (Pang et al., 2002) is a simple probabilistic classifier based on Bayes' theorem and is particularly suited when the dimensionality of the inputs are high. NB classification is an approach to text classification that assigns the class;

$$c^{*} = \arg\max_{c} P(c \mid d)$$

to a given document d. Its underlying probability model can be described as an 'independent feature model'. The *NB* classifier uses the Bayes' rule,

$$P(c \mid d) = \frac{P(c) * P(d \mid c)}{P(d)}$$

where $P(d)$ plays no role in selecting c^{*}. To estimate the term $P(d \mid c)$, NB decomposes it by assuming the f_i's are conditionally independent given d's class as in (2).

$$P_{\mathrm{NB}}(c \mid d) := \frac{P(c)\left(\prod^{m}{}_{i=1} P\left(f_i \mid c\right)^{ni(d)}\right)}{P(d)}$$

where m is the number of features and f_i is the feature vector.

Consider a training method consisting of a relative-frequency estimation $P(c)$ and $P(f_i \mid c)$. Despite its simplicity and the fact that its conditional independence assumption clearly does not hold in real-world situations, NB-based text categorisation still tends to perform surprisingly well; indeed, NB is optimal for certain problem classes with highly dependent features.

4.2 Classification using maximum entropy

Maximum entropy classification (Pang et al., 2002) (MaxEnt, or ME, for short) is an alternative technique which has proven effective in a number of natural language processing applications (Berger et al., 1996). Nigam et al. (1999) show that it sometimes, but not always, outperforms NB at standard text classification. Its estimate of $P(c \mid d)$ takes the following exponential form.

4.3 Classification using support vector

SVMs have been shown to be highly effective at traditional text categorisation, generally outperforming NB. They are *large-margin*, rather than probabilistic, classifiers, in contrast to NB and MaxEnt (Pang et al., 2002). The basic idea behind SVMs (Yaquan and Haibo, 2011) is to find a separating hyperplane with the largest margin in a given higher-dimensional feature space. The search for this hyperplane corresponds to a constrained optimisation problem.

4.4 Classification using score function

A custom score function for review sentiment classification was given by Dave et al. The algorithm consist of two steps:

Step 1 It scores each term in the training set using the following equation,

$$score(t_i) = \frac{P_r(t_i \mid C) - P_r(t_i \mid C')}{P_r(t_i \mid C) + P_r(t_i \mid C')}$$

where t_i is a term and C is a class and C' is its complement, i.e., not C, and $Pr(t_i \mid C)$ is the conditional probability of term t_i in class C. It is computed by taking the number of times that a term t_i occurs in class C reviews and dividing it by the total number of terms in the reviews of class C. A term's score is thus a measure of bias towards either class ranging from –1 and 1.

Step 2 To classify a new document $d_i = t_1 \ldots t_n$, the algorithm sums up the scores of all terms and uses the sign of the total to determine the class. That is, it uses the following equation for classification,

$$class(d_i) = \begin{cases} C & eval(d_i) > 0 \\ C' & \text{otherwise,} \end{cases}$$

where $eval(d_i) = \sum_j score\ (t_j)$.

Experiments were conducted based on a large number of reviews (more than 3,000) of seven types of products. The results showed that the bigrams (consecutive two words) and trigrams (consecutive three words) as terms gave (similar) best accuracies (84.6%–8.3%), on two different review datasets. No stemming or stopword removal was applied.

5 Aspect-based semantic classification at document level

This task treat opinion mining as text classification problem. It classifies evaluative text as being positive or negative. For instance, given a product review, the system determines whether the reviewer's attitude is positive or negative. This classification is typically at document-level. Given a set of evaluative texts D, a sentiment classifier classifies each document, d € D into one of the two classes positive and negative. Positive means that the document d expresses positive opinion. Negative means that the document d expresses negative opinion. For example, given a smart-phone review the classifier classifies the review either as positive review or negative review.

5.1 Opinion extraction

The fundamental step of opinion mining is to extract the reviews depending upon its subjectivity and relevance. These reviews are extracted based on positive and negative sentiment words and phrases contained in each evaluative text. The extraction is based on concept of natural language processing technique, given by Turney (2002). This technique is called as POS tagging. POS tagging is the task of labelling (or tagging) each word in a sentence with its appropriate part of speech. The POS of a word is a linguistic category that is defined by its syntactic or morphological behaviour. The common POS categories in English grammar are: nouns, adjectives, verbs, adverb, pronoun, preposition, conjunction and interjection. Also, there exist many categories which arise from different forms of these categories. For example, a verb can be a verb in its base form, in its past tense, etc. The standard Penn Treebank POS tags (shown in Table 1) have been used for tagging the words of the reviews.

Table 1 Penn Treebank POS tags

Tag	*Description*	*Tag*	*Description*
CC	Coordinating conjunction	PRP$	Possessive pronoun
CD	Cardial number	RB	Adverb
DT	Determiner	RBR	Adverb, comparative
EX	Existential *there*	RBS	Adverb, superlative
FW	Foreign word	RP	Particle
IN	Preposition or subordinating conjunction	SYM	Symbol
JJ	Adjective	TO	To
JJR	Adjective, Comparative	UH	Interjection
JJS	Adjective, superlative	VB	Verb, base form
LS	List item marker	VBD	Verb, past tense
MD	Modal	VBG	Verb, gerund or present particle
NN	Noun, singular or mass	VBN	Verb, past particle
NNS	Noun, plural	VBP	Verb, non-3rd person singular present
NNP	Proper noun, singular	VBZ	Verb, 3rd person singular present
NNPS	Proper noun, plural	WDT	Wh-determiner
PDT	Predeterminer	WP	Wh-pronoun
POS	Possessive ending	WP$	Possessive wh-pronoun
PRP	Personal pronoun	WRB	Wh-adverb

The extracted words are basically adjectives or adverbs as they are good indicators of subjectivity and opinions. However, although an isolated adjective may indicate subjectivity, there may be an insufficient context to determine its opinion orientation. For example, the adjective ‘unpredictable’ may have a negative orientation in an automotive review, in such a phrase as ‘unpredictable steering’, but it could have a positive orientation in a movie review, in a phrase such as 'unpredictable plot'. Therefore, extracting two consecutive words, where one member of the pair is an adjective/adverb and the other is a context word is necessary. For example, the pattern in line 2 of Table 2 means that two consecutive words are extracted if the first word is an adverb and the second word is an adjective, but the third word (which is not extracted) cannot be a noun. NNP and NNPS are avoided so that the names of the objects in the review cannot influence the classification.

Table 2 Patterns of tags for extracting two-words phrases from reviews

S. no	*First word*	*Second word*	*Third word (not extracted)*
1	JJ	NN or NNS	Anything
2	RB, RBR, or RBS	JJ	Not NN nor NNS
3	JJ	JJ	Not NN nor NNS
4	NN or NNS	JJ	Not NN nor NNS
5	RB, RBR, or RBS	VB, VBD, VBN, VBG	Anything

Taking review *this phone takes amazing pictures* as an example, the extracted phrase would be *amazing picture* as it satisfies the first pattern.

Opinion extraction using NB: Example.

Consider the tagged review of a smartphone in Figure 4.

Figure 4 Tagged review

Nice_JJ looking_VBG design_NN ,_, looks_VBZ to_TO have_VB decent_JJ battery_NN life_NN ,_, light_JJ weight_NN and_CC decent_JJ speed_NN on_IN wifi_NNS and_CC LTE_NN A_DT good_JJ improvement_NN over_IN the_DT iPhone_NNP 4_CD but_CC not_RB the_DT 4S_NN

The extracted phrases by using tagging rules are: decent battery, light weight, good improvement

The opinion classification of the above review is positive by using Bayes formula. This is explained as follows. The probability of the document belonging to class positive is hundred percent whereas the probability of the document belonging to class negative is null. Hence the above review is classified as positive.

6 Comparative and feature based opinion mining

6.1 Feature-based opinion mining

Opinion mining can be done on sentence level, document level and feature level. A positive evaluative text does not mean that author has positive opinion on every aspect of that product. Similarly if negative evaluative text is their then it does not mean that their can not be some positive aspect in the text. Therefore feature-based opinion mining is the another method which can be used to consider both positive and negative feature in a review. Two main task which are apparent in feature-based mining are.

We have to identify and extract those feature of product for which reviewer has expressed their opinion. Determining whether the opinions on feature are positive, negative or neutral. An opinion can be expressed on a product, an individual, an organisation, an event or any other topic. So in general it can be treated as a object which may have sub components.

Example 1: A particular brand of phone is an object, it has set of component, e.g., RAM, operating system and its version, speed, picture quality, battery life, screen size, etc.

According to the definition given by Liu (2011),

Definition (object): An object O is an entity which can be a product, person, event, organisation, or topic. It is associated with a pair, O: (T, A), where T is a hierarchy or taxonomy of components (or parts), subcomponents, and so on, and A is a set of attributes of O. Each component has its own set of sub-components and attributes.

A feature can be explicit feature or implicit feature. Let the evaluative text (e.g., a product review) be r. In the most general case, r consists of a sequence of sentences $r = <s_1, s_2, \ldots, s_m>$ (Liu, 2011).

Definition (explicit and implicit feature): If a feature f appears in evaluative text r, it is called an explicit feature in r. If f does not appear in r but is implied, it is called an implicit feature in r.

Definition (explicit and implicit opinion): An *explicit opinion* on feature f is a subjective sentence that directly expresses a positive or negative opinion. An *implicit opinion* on feature f is an objective sentence that implies a positive or negative opinion (Liu, 2011).

Example 2: The following sentence expresses an explicit positive opinion:

> "The picture quality of this Phone is amazing".

The following sentence expresses an implicit negative opinion:

> "The earphone broke in two days".

The final output for each evaluative text d is a set of pairs. Each pair is denoted by (f, SO), where f is a feature and SO is the semantic or opinion orientation (positive or negative) expressed in d on feature f. We ignore neutral opinions in the output as they are not usually useful. Below is the example of a feature-based summary of opinions

*Mobile phone*_1:

Feature: *picture quality*

Positive: 133 <individual review sentences>

Negative: 8 <individual review sentences>

Feature: *size*

Positive: 92 <individual review sentences>

Negative: 12 <individual review sentences>

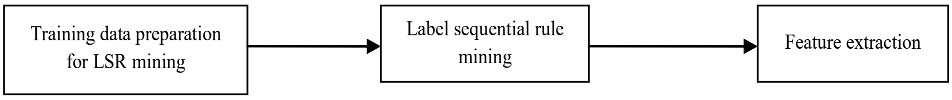
6.1.1 Object feature extraction

Feature extraction work is mainly carried out from online product reviews, there could be different review formats available online.

Format 1: pros, cons and the detailed review

In this format reviewer is asked to describe pros and cons of the product separately, e.g., a product feature can be expressed with a noun, adjective, verb or adverb. The labels and their POS tags used in mining LSRs are: {$feature, NN}, {$feature, JJ}, {$feature, VB} and {$feature, RB}, where $feature denotes a feature to be extracted, and NN stands for noun, VB for verb, JJ for adjective, and RB for adverb. They represent both explicit features and implicit feature indicators. We call a word that indicates an implicit feature an *implicit feature indicator*. For example, in the sentence 'this camera is too heavy', 'heavy' is an adjective and is an implicit feature indicator for feature 'weight' (Liu, 2011). Given a set of reviews, this method consist following steps.

Figure 5 Various steps in feature-based mining



Feature extraction from other formats

"It is a great fast processing smartphone for this century"

September 1, 2012.

Pros:

It is large size, and the rotatable lens is great. It is very easy to use, and has fast response from the shutter. The LCD …

Cons:

It almost has no cons. It could be better if the LCD is bigger and it is going to be best if the model is designed to a smaller size.

Figure 6 Review formats of different types

I did a lot of research last year before I bought this smart phone... It kinda hurt to leave behind my beloved Nikon 35 mm SLR, but I was going to Italy, and I needed something smaller, and digital.
The pictures coming out of this phone are amazing. The 'auto' feature takes great pictures most of the time. And with digital, you're not wasting film if the picture doesn't come out. …

Below, we describe an unsupervised method for finding explicit features that are nouns and noun phrases. This method requires a large number of reviews, and consists of two steps:

1 Finding frequent nouns and noun phrases. Nouns and noun phrases (or groups) are identified by using a POS tagger. We then count their frequency and only keep the frequent ones. A frequency threshold can be decided experimentally. The reason for using this approach is that most product features are nouns, and those nouns that are frequently talked about are usually genuine and important features. Irrelevant contents in reviews are often diverse, i.e., they are quite different in different reviews. When people comment on product features, the vocabulary that they use converges. Those nouns that are infrequent are likely to be non-features or less important features.

2 Finding infrequent features by making use of sentiment words. Sentiment words (also called opinion words) are usually adjectives and adverbs that express positive or negative opinions, e.g., great, amazing, bad, and expensive. The idea is as follows: The same opinion word can be used to describe different objects. Opinion words that modify frequent features can be used to find infrequent features (Liu, 2011).

6.2 Comparative sentence and relation mining

In the earlier section we have explored one form of evaluation which is based on positive and negative opinions expressed by the reviewer. Comparing the object with another object in a product review is another way in opinion mining. Most of the customer who wish to purchase some new product does the comparative analysis which is one of the most prominent way to know about the product. Comparative analysis gives better opinion about the product as its feature are being compared with other similar category product. For example, "the battery life of Moto G is better than Samsung Grand".

A comparison may be subjective or objective depending how it is expressed above example is subjective type comparison. On the other hand objective type comparison is done on some quantitative values about the product, e.g., "the screen size of Moto G is 4.7, while that of Samsung Quattro is 5.0".

A *comparative sentence* is a sentence that expresses a relation based on similarities or differences of more than one object. The comparison in a comparative sentence is usually expressed using the *comparative* or the *superlative* form of an adjective or adverb.

Types of comparison: Comparison can be classified into four main types, three types of comparison are gradable comparison while fourth one is non-gradable comparison. The gradable comparison are based on the relationships of less than or greater, equal and greater or less than all others.

1 Non-equal gradable comparison: in this type of comparison relation of the type greater or less than express an ordering of same objects with regard to their feature, e.g., "the capacitive touch is better than resistive touch of phone".

2 Equalative comparison: in this type of comparison the state of object are equal with respect to some of their fact, e.g., "the picture quality of phone A is as good as of phone B".

3 Superlative comparison: relations of the type greater or less than all *others* that rank one object over *all* others, e.g., "the Android kitkat is the fastest".

4 Non-gradable comparisons: sentences that compare features of two or more objects, but do not grade them. There are three main types:

- object A is similar to or different from object B with regard to some features, e.g., “Sprite tastes differently from Mountain dew”
- object A has feature $f1$, and object B has feature $f2$ ($f1$ and $f2$ are usually substitutable), e.g., “desktop PCs use external speakers but laptops use internal speakers”
- object A has feature f, but object B does not have, e.g., “cell phone A has rear camera, but cell phone B does not have”.

Given an evaluative text d, *comparison mining* consists of two tasks:

1 Identify comparative passages or sentences from d, and classify the identified comparative sentences into different types or classes.

2 Extract comparative relations from the identified sentences. This involves the extraction of entities and their features that are being compared, and the comparative keywords. Relations in gradable adjectival comparisons can be expressed with

$$\left(< relationWord >,\ < features >,\ < entityS1 >,\ < entityS2 >,\ < type >\right)$$

where

relationWord: the comparative keyword used to express a comparative relation in a sentence.

features: a set of features being compared.

*entityS*1 and *entityS*2: sets of entities being compared. Entities in *entityS*1 appear to the left of the relation word and entities in *entityS*2 appear to the right of the relation word.

type: non-equal gradable, equative or superlative (Liu, 2011).

7 Conclusions

In this paper we have described about sentiment analysis and the work which has been carried out by several researcher in this area. So we have seen that research work in this particular area has been geared up, researcher have found many useful techniques through which the data, which is in the form of text (sentence, document) can be mine effectively for analysing people’s views. Opinion mining and subjectivity analysis have given us opportunity to mine opinions and reviews which can be found on blogs, social network, twitter, and various product reviews websites. We have also explored in detail the various techniques which are used in sentiment classification, and presents a brief comparison between various classification techniques. We have also presented that how various opinion can be extracted from reviews, to explain this further we have also shown some example that how we can extract information from the text using various classification techniques. We have also discussed that how feature-based mining and comparative sentence mining can be useful for better product analysis where different types of product can be analysed on the basis of their feature and also in comparison to other products. Our approach to opinion mining takes inspiration from a number of sources. It is most similar to the work of Turney (2002) in terms of technique. The approaches in sentiment analysis need to employ some more efficient strategies to deal with the linguistic issues imposed.

References

Akkaya, C., Wiebe, J. and Mihalcea, R. (2009) 'Subjectivity word sense disambiguation', in *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing (EMNLP-2009)*.

Argamon-Engelson, S., Koppel, M. and Avneri, G. (1998) 'Style-based text categorization: What newspaper am I reading?', *Proc. of the AAAI Workshop on Text Categorization*, pp.1–4.

Berger, A., Della, S., Pietra and Della, V. (1996) 'A maximum-entropy approach to natural language processing', *Computational Linguistics*, March, Vol. 22, No. 1.

Bollen, J. and Mao, H. (2011) 'Twitter mood as a stock market predictor', *IEEE Computer*, October, Vol. 44, No. 10, pp.91–94.

Dellarocas, C., Zhang, X.M. and Awad, N.F. (2007) 'Exploring the value of online product reviews in forecasting sales: the case of motion pictures', *Journal of Interactive Marketing*, Vol. 21, No. 4, pp.23–45.

Devitt, A. and Ahmad, K. (2007) 'Sentiment analysis in financial news: a cohesion based approach', *Proceedings of the Association for Computational Linguistics (ACL)*, pp.984–991.

Dimitrova, M., Finn, A., Kushmerick, N. and Smyth, B. (2002) 'Web genre visualisation', *Conference on Human Factors in Computing Systems*.

Dini L. and Mazzini, G. (2002) 'Opinion classification through information extraction', *Proceedings of the Conference on Data Mining Methods and Databases for Engineering, Finance and Other Fields (Data Mining)*, pp.299–310.

Durbin, S.D., Neal Richter, J. and Warner, D. (2003) 'A system for affective rating of texts', *Proceedings of OTC-03, 3rd Workshop on Operational Text Classification*, Washington, USA.

Efron, M. (2004) 'Cultural orientation: classifying subjective documents by cociation [sic] analysis', *Proceedings of the AAAI Fall Symposium on Style and Meaning in Language, Art, Music, and Design*, pp.41–48.

Esuli, A. and Sebastiani, F. (2006) 'Determining term subjectivity and term orientation for opinion mining', *Proceedings of the European Chapter of the Association for Computational Linguistics (EACL)*.

Gamon, M., Aue, A., Corston-Oliver, S. and Ringger, E. (2005) 'Pulse: mining customer opinions from free text', *Advances in Intelligent Data Analysis VI*, pp.121–132.

Glance, N.S., Hurst, M. and Tomokiyo, T. (2004) 'Blogpulse: automated trend discovery for weblogs', *WWW 2004 Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics*.

Go, A., Bhayani, R. and Huang, L. (2009) *Twitter Sentiment Classification Using Distant Supervision*, Technical report, Stanford.

Grefenstette, G., Qu, Y., Shanahan, J.G. and Evans, D.A. (2004) 'Coupling niche browsers and affect analysis for an opinion mining application', *Proceedings of Recherche d'Information Assist´ee par Ordinateur (RIAO)*.

Hassan, A. and Radev, D. (2010) 'Identifying text polarity using random walks', in *Proceedings of Annual Meeting of the Association for Computational Linguistics (ACL-2010)*.

Hatzivassiloglou, V. and McKeown, K. (1997) 'Predicting the semantic orientation of adjectives', in *Proceedings of Annual Meeting of the Association for Computational Linguistics (ACL-1997)*.

Hatzivassiloglou, V. and Wiebe, J. (2000) 'Effects of adjective orientation and gradability on sentence subjectivity', in *Proceedings of International Conference on Computational Linguistics (COLING-2000)*.

He, Y., Lin, C. and Alani, H. (2011) 'Automatically extracting polarity-bearing topics for cross-domain sentiment classification', in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics (ACL-2011)*.

Hillard, D., Ostendorf, M. and Shriberg, E. (2003) 'Detection of agreement vs. disagreement in meetings: training with unlabeled data', *Proceedings of HLTNAACL*.

Hu, M. and Liu, B. (2004) 'Mining and summarizing customer reviews', in *Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004)*.

Indurkhya, N. and Damerau, F. (2010) *Handbook of Natural Language Processing,* Chapman & Hall/CRC Machine Learning & Pattern Recognition, Chapman & Hall/CRC Taylor & Francis Group, Boca Raton, FL.

Jijkoun, V., Rijke, M.d. and Weerkamp, W. (2010) 'Generating focused topic-specific sentiment Lexicons', in *Proceedings of Annual Meeting of the Association for Computational Linguistics (ACL-2010)*.

Jindal, N. and Liu, B. (2006a) 'Identifying comparative sentences in text documents', in *Proceedings of ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR-2006)*.

Jindal, N. and Liu, B. (2006b) 'Mining comparative sentences and relations', in *Proceedings of National Conf. on Artificial Intelligence (AAAI-2006)*.

Kessler, B., Nunberg, G. and Schütze, H. (1997) 'Automatic detection of text genre', *Proc. of the 35th ACL/8th EACL*, pp.32–38.

Kobayashi, N., Inui, K., Tateishi, K. and Fukushima, T. (2004) 'Collecting evaluative expressions for opinion extraction', *IJCNLP*, pp.596–605.

Li, S., Lee, S.Y.M., Chen, Y., Huang, C.R. and Zhou, G. (2010a) 'Sentiment classification and polarity shifting', in *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*.

Li, S., Lin, C., Song, Y. and Li, Z. (2010b) 'Comparable entity mining from comparative questions', in *Proceedings of Annual Meeting of the Association for Computational Linguistics*
(ACL-2010).

Lin, W-H., Wilson, T., Wiebe, J. and Hauptmann, A. (2006) 'Which side are you on? Identifying perspectives at the document and sentence levels', *Proceedings of the Conference on Natural Language Learning (CoNLL)*.

Liu, B. (2010) 'Sentiment analysis and subjectivity', in N. Indurkhya and F.J. Damerau (Eds.): *Handbook of Natural Language Processing*, 2nd ed.

Liu, B. (2011) *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*, 1st ed. (2006) and 2nd ed. (2011), Springer, Verlag Berlin Heidelberg 2007.

Martin, J.R. and White, P.R.R. (2005) *The Language of Evaluation: Appraisal in English*, Palgrave, London, UK.

Meng, X. and Wang, H. (2009) 'Mining user reviews: from specification to summarization', in *Proceedings of the ACL-IJCNLP 2009 Conference* Short Papers.

Moghaddam, S. and Popowich, F. (2010) *Opinion Polarity Identification through Adjectives*, CoRR abs/1011.4623.

Mullaly, A.C., Gagné, C.L., Spalding, T.L. and Marchak, K.A. (2010) 'Examining ambiguous adjectives in adjective-noun phrases: evidence for representation as a shared core-meaning with sense specialization', *The Mental Lexicon* Vol. 5, No. 1, pp.87–114, DOI: 10.1075/ml.5.1.04mul.

Mullen, T. and Collier, N. (2004) 'Sentiment analysis using support vector machines with diverse information sources', in *Proceedings of EMNLP-2004*.

Nigam, K., Lafferty, J. and McCallum, A. (1999) 'Using maximum entropy for text classification', *Proc. of the IJCAI-99 Workshop on Machine Learning for Information Filtering*, pp.61–67, Ted Pedersen, 2001.

Osgood, C.E., Suci, G.J. and Tannenbaum, P.H. (1967) *The Measurement of Meaning*, University of Illinois Press.

Ounis, I., Macdonald, C. and Soboroff, I. (2008) 'On the TREC blog track', *Proceedings of the International Conference on Weblogs and Social Media (ICWSM)*.

Pak, A. and Paroubek, P. (2010) 'Twitter as a corpus for sentiment analysis and opinion mining', *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, May, European Language Resources Association (ELRA), Valletta, Malta.

Pang, B. and Lee, L. (2004) 'A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts', in *Proceedings of Meeting of the Association for Computational Linguistics (ACL-2004)*.

Pang, B. and Lee, L. (2008) 'Opinion mining and sentiment analysis', *Foundations and Trends in Information Retrieval*, Vol. 2, Nos. 1–2, pp.1–135.

Pang, B., Lee, L. and Vaithyanathan, S. (2002) 'Thumbs up?: Sentiment classification using machine learning techniques', in *Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP-2002)*.

Popescu, A. and Etzioni, O. (2005) 'Extracting product features and opinions from reviews', in *Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP-2005)*.

Prabowo, R. and Thelwall, M. (2009) 'Sentiment analysis: a combined approach', *Journal of Informetrics*, Vol. 3, No. 2, pp.143–157, doi:10.1016/j.joi.2009.01.003.

Rauber, A. and Muller-Kogler, A. (2001) 'Integrating automatic genre analysis into digital libraries', *First ACM-IEEE Joint Conference on Digital Libraries*.

Santorini, B. (1990) *Part-of-speech Tagging Guidelines for the Penn Treebank Project*, University of Pennsylvania, School of Engineering and Applied Science, Dept. of Computer and Information Science, Philadelphia.

Scharl, A. and Weichselbraun, A. (2008) 'An automated approach to investigating the online media coverage of US presidential elections', *Journal of Information Technology & Politics*, Vol. 5, No. 1, pp.121–132.

Shanahan, J., Qu, Y. and Wiebe, J. (2006) *Computing Attitude and Affect in Text: Theory and Applications*, Springer-Verlag, Heidelberg.

Spertus, E. (1997) 'Smokey: automatic recognition of hostile messages', *Proc. of Innovative Applications of Artificial Intelligence*, IAAI, pp.1058–1065

Subasic, P. and Huettner, A. (2001) 'Affect analysis of text using fuzzy semantic typing', *IEEE Transactions on Fuzzy Systems*, Vol. 9, pp.483–496.

Sugato, C., Sabyasachi, B. and Karishma, B. (2013) 'On evaluation of customer dissatisfaction index based on sentiment analysis: a statistical approach', *Int. J. of Data Analysis Techniques and Strategies*, Vol. 5, No. 4, pp.427–444.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K. and Stede, M. (2011) 'Lexicon-based methods for sentiment analysis', *Computational Linguistics*, Vol. 37, pp.267–307.

Thomas, M., Pang, B. and Lee, L. (2006) 'Get out the vote: determining support or opposition from congressional floor debate transcripts', *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP 2006)*, Sydney, pp.327–335.

Tong, R. (2001) 'An operational system for detecting and tracking opinions in on-line discussion', *SIGIR 2001 Workshop on Operational Text Classification*.

Turney, P. (2002) 'Thumbs up or thumbs down?: Semantic orientation applied to unsupervised classification of reviews', in *Proceedings of Annual Meeting of the Association for Computational Linguistics (ACL-2002)*.

Weichselbraun, A., Gindl S. and Scharl, A. (2010) 'A context-dependent supervised learning approach to sentiment detection in large textual databases', *Journal of Information and Data Management*, October, Vol. 1, No. 3, pp.329–342.

Whitelaw, C., Garg, N. and Argamon, S. (2005) 'Using appraisal groups for sentiment analysis,' *Proceedings of ACM SIGIR Conference on Information and Knowledge Management (CIKM 2005)*, pp.625–631.

Wiebe, J. and Riloff, E. (2005) 'Creating subjective and objective sentence classifiers from unannotated texts', *Computational Linguistics and Intelligent Text Processing*, Vol. 3406, pp.486–497, Hedilerberg Springer, DOI: 10.1007/978-3-540-30586-6_53.

Wilson, T., Wiebe, J. and Hoffmann, P. (2005) 'Recognizing contextual polarity in phrase-level sentiment analysis', *Proceedings of the Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP)*, pp.347–354.

Wilson, T., Wiebe, J. and Hoffmann, P. (2009) 'Recognizing contextual polarity: an exploration of features for phrase-level sentiment analysis', *Computational Linguistics*, Vol. 35, No. 3, 399–433.

Witten, I.H. and Frank, E. (1999) *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations* (The Morgan Kaufmann Series in Data Management Systems), 1st ed., Morgan Kaufmann.

Yaquan, X. and Haibo, W. (2011) 'A new feature selection method based on support vector machines for text categorization', *Int. J. of Data Analysis Techniques and Strategies*, Vol. 3, No. 1, pp.1–20.

Yugowati, P., Shaou-Gang, M. and Hui-Ming, W. (2013) 'Supervised learning approaches and feature selection – a case study in diabetes', *Int. J. of Data Analysis Techniques and Strategies*, Vol. 5, No. 3, pp.323–337.

Zaidan, O.F., Eisner, J. and Piatko, C.D. (2007) 'Using annotator rationales to improve machine learning for text categorization', *Proceedings of NAACL HLT*, pp.260–267.