Sentiment Analysis and Opinion Mining: A Survey

Article · June 2012

2 authors:

Vinodhini G
Annamalai University

Dr. RM Chandrasekaran
Annamalai University

Volume 2, Issue 6, June 2012 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Sentiment Analysis and Opinion Mining: A Survey

G.Vinodhini*
Assistant Professor, Department of Computer Science and Engineering, Annamalai University, Annamalai Nagar-608002.
India

RM.Chandrasekaran
Professor, Department of Computer Science and Engineering, Annamalai University, Annamalai Nagar-608002.
India

ABSTRACT: **Due to the sheer volume of opinion rich web resources such as discussion forum, review sites , blogs and news corpora available in digital form, much of the current research is focusing on the area of sentiment analysis. People are intended to develop a system that can identify and classify opinion or sentiment as represented in an electronic text. An accurate method for predicting sentiments could enable us, to extract opinions from the internet and predict online customer's preferences, which could prove valuable for economic or marketing research. Till now, there are few different problems predominating in this research community, namely, sentiment classification, feature based classification and handling negations. This paper presents a survey covering the techniques and methods in sentiment analysis and challenges appear in the field.**

Key words: **sentiment, opinion, machine learning, semantic.**

1. INTRODUCTION

Sentiment analysis is a type of natural language processing for tracking the mood of the public about a particular product or topic. Sentiment analysis, which is also called opinion mining, involves in building a system to collect and examine opinions about the product made in blog posts, comments, reviews or tweets. Sentiment analysis can be useful in several ways. For example, in marketing it helps injudging the success of an ad campaign or new product launch, determine which versions of a product or service are popular and even identify which demographics like or dislike particular features.

There are several challenges in Sentiment analysis. The first is a opinion word that is considered to be positive in one situation may be considered negative in another situation. A second challenge is that people don't always express opinions in a same way. Most traditional text processing relies on the fact that small differences between two pieces of text don't change the meaning very much. In Sentiment analysis, however, "the picture was great" is very different from "the picture was not great". People can be contradictory in their statements. Most reviews will have both positive and negative comments, which is somewhat manageable by analyzing sentences one at a time. However, in the more informal medium like twitter or blogs, the more likely people are to combine different opinions in the same sentence which is easy for a human to understand, but more difficult for a computer to parse. Sometimes even other people have difficulty understanding what someone thought based on a short piece of text because it lacks context. For example, "That movie was as good as its last movie" is entirely dependent on what the person expressing the opinion thought of the previous model.

The user's hunger is on for and dependence upon online advice and recommendations the data reveals is merely one reason behind the emerge of interest in new systems that deal directly with opinions as a first-class object. Sentiment analysis concentrates on attitudes, whereas traditional text mining focuses on the analysis of facts. There are few main fields of research predominate in Sentiment analysis: sentiment classification, feature based Sentiment classification and opinion summarization. Sentiment classification deals with classifying entire documents according to the opinions towards certain objects. Feature-based Sentiment classification on the other hand considers the opinions on features of certain objects. Opinion summarization task is different from traditional text summarization because only the features of the product are mined on which the customers have expressed their opinions. Opinion summarization does not summarize the reviews by selecting a subset or rewrite some of the original sentences from the reviews to capture the main points as in the classic text summarization.

Languages that have been studied mostly are English and in Chinese .Presently, there are very few researches conducted on sentiment classification for other languages like Arabic, Italian and Thai. This survey aims at focusing much of the work in English and a few from Chinese. The emergence of sentiment analysis dates back to late 1990's, but becomes a major emerging sub field of information management discipline only from 2000, especially from 2004 onwards, which this survey focuses.

For the sake of convenience the remainder of this paper is organized as follows: Section 2 presents the data sources used for opinion mining. Section 3 introduces machine learning and semantic orientation approaches for sentiment classification. Section 4 presents some applications of sentiment classification. Then we present some tools available for sentiment classification in section 4. The fifth section is about the performance evaluation done. Last section concludes our study and discusses some future directions for research.

2. DATA SOURCE

User's opinion is a major criterion for the improvement of the quality of services rendered and enhancement of the deliverables. Blogs, review sites, data and micro blogs provide a good understanding of the reception level of the products and services.

2.1. Blogs

With an increasing usage of the internet, blogging and blog pages are growing rapidly. Blog pages have become the most popular means to express one's personal opinions. Bloggers record the daily events in their lives and express their opinions, feelings, and emotions in a blog (Chau & Xu, 2007). Many of these blogs contain reviews on many products, issues, etc. Blogs are used as a source of opinion in many of the studies related to sentiment analysis (Martin, 2005; Murphy, 2006; Tang et al., 2009).

2.2. Review sites

For any user in making a purchasing decision, the opinions of others can be an important factor. A large and growing body of user-generated reviews is available on the Internet. The reviews for products or services are usually based on opinions expressed in much unstructured format. The reviewer's data used in most of the sentiment classification studies are collected from the e-commerce websites like www.amazon.com (product reviews), www.yelp.com (restaurant reviews), www.CNET download.com (product reviews) and www.reviewcentre.com, which hosts millions of product reviews by consumers. Other than these the available are professional review sites such as www.dpreview.com , www.zdnet.com and consumer opinion sites on broad topics and products such as www .consumerreview.com, www.epinions.com, www.bizrate.com (Popescu& Etzioni ,2005 ; Hu,B.Liu ,2006 ; Qinliang Mia, 2009; Gamgaran Somprasertsi ,2010).

2.3. DataSet

Most of the work in the field uses movie reviews data for classification. Movie review datas are available as dataset (http:// www.cs.cornell.edu/People/pabo/movie-review-data). Other dataset which is available online is multi-domain sentiment (MDS) dataset. (http:// www.cs.jhu.edu/mdredze/datasets/sentiment). The MDS dataset contains four different types of product reviews extracted from Amazon.com including Books, DVDs, Electronics and Kitchen appliances, with 1000 positive and 1000 negative reviews for each domain. Another review dataset available is http://www.cs.uic.edu/liub/FBS/CustomerReviewData.zip. This dataset consists of reviews of five electronics products downloaded from Amazon and Cnet (Hu and Liu ,2006; Konig & Brill ,2006 ; Long Sheng ,2011; Zhu Jian ,2010 ; Pang and Lee ,2004; Bai et al. ,2005; Kennedy and Inkpen ,2006; Zhou and Chaovalit ,2008; Yulan He 2010; Rudy Prabowo ,2009; Rui Xia ,2011).

2.4. Micro-blogging

Twitter is a popular microblogging service where users create status messages called "tweets". These tweets sometimes express opinions about different topics. Twitter messages are also used as data source for classifying sentiment.

3. Sentiment Classification

Much research exists on sentiment analysis of user opinion data, which mainly judges the polarities of user reviews. In these studies, sentiment analysis is often conducted at one of the three levels: the document level, sentence level, or attribute level. In relation to sentiment analysis, the literature survey done indicates two types of techniques including machine learning and semantic orientation. In addition to that, the nature language processing techniques (NLP) is used in this area, especially in the document sentiment detection. Current-day sentiment detection is thus a discipline at the crossroads of NLP and Information retrieval, and as such it shares a number of characteristics with other tasks such as information extraction and text-mining, computational linguistics, psychology and predicative analysis.

3.1. Machine Learning

The machine learning approach applicable to sentiment analysis mostly belongs to supervised classification in general and text classification techniques in particular. Thus, it is called ''supervised learning". In a machine learning based classification, two sets of documents are required: training and a test set. A training set is used by an automatic classifier to learn the differentiating characteristics of documents, and a test set is used to validate the performance of the automatic classifier. A number of machine learning techniques have been adopted to classify the reviews. Machine learning techniques like Naive Bayes (NB), maximum entropy (ME), and support vector machines (SVM) have achieved great success in text categorization. The other most well-known machine learning methods in the natural language processing area are

K-Nearest neighbourhood, ID3, C5, centroid classifier, winnow classifier, and the N-gram model.

Naive Bayes is a simple but effective classification algorithm. The Naive Bayes algorithm is widely used algorithm for document classification (Melville et al., 2009; Rui Xia, 2011; Ziqiong, 2011; Songho tan, 2008 and Qiang Ye, 2009). The basic idea is to estimate the probabilities of categories given a test document by using the joint probabilities of words and categories. The naive part of such a model is the assumption of word independence. The simplicity of this assumption makes the computation of Naive Bayes classifier far more efficient.

Support vector machines (SVM), a discriminative classifier is considered the best text classification method (Rui Xia, 2011; Ziqiong, 2011; Songho tan, 2008 and Rudy Prabowo, 2009). . The support vector machine is a statistical classification method proposed by Vapnik . Based on the structural risk minimization principle from the computational learning theory, SVM seeks a decision surface to separate the training data points into two classes and makes decisions based on the support vectors that are selected as the only effective elements in the training set. Multiple variants of SVM have been developed in which Multi class SVM is used for Sentiment classification (Kaiquan Xu, 2011).

The idea behind the centroid classification algorithm is extremely simple and straightforward (Songho tan, 2008). Initially the prototype vector or centroid vector for each training class is calculated, then the similarity between a testing document to all centroid is computed, finally based on these similarities, document is assigned to the class corresponding to the most similar centroid.

The K-nearest neighbor (KNN) is a typical example based classifier that does not build an explicit, declarative representation of the category, but relies on the category labels attached to the training documents similar to the test document. Given a test document d, the system finds the k nearest neighbors among training documents. The similarity score of each nearest neighbor document to the test document is used as the weight of the classes of the neighbor document (Songho tan, 2008).

Winnow is a well-known online mistaken-driven method. It works by updating its weights in a sequence of trials. On each trial, it first makes a prediction for one document and then receives feedback; if a mistake is made, it updates its weight vector using the document. During the training phase, with a collection of training data, this process is repeated several times by iterating on the data (Songho tan, 2008). Besides these classifiers other classifiers like ID3 and C5 are also investigated (Rudy Prabowo, 2009).

Besides using these above said machine learning methods individually for sentiment classification, various comparative studies have been done to find the best choice of machine learning method for sentiment classification. Songbo Tan (2008) presents an empirical study of sentiment categorization on Chinese documents. He investigated four feature selection methods (MI,IG, CHI and DF) and five learning methods (centroid classifier, K-nearest neighbor, winnow classifier, Naive Bayes and SVM) on a Chinese sentiment corpus. From the results he concludes that, IG performs the best for sentimental terms selection and SVM exhibits the best performance for sentiment classification. When applying SVM, naive Bayes and n-gram model to the destination reviews, Ye et al. (2009) found that SVM outperforms the other two classifiers.

Rudy Prabowo (2009) described an extension by combining rule-based classification, supervised learning and machine learning into a new combined method. For each sample set, they carried out 10-fold cross validation. For each fold, the associated samples were divided into training and a test set. For each test sample, a hybrid classification is carried out, i.e., if one classifier fails to classify a document, the classifier passes the document onto the next classifier, until the document is classified or no other classifier exists. Given a training set, the Rule Based Classifier (RBC) used a Rule Generator to generate a set of rules and a set of antecedents to represent the test sample and used the rule set derived from the training set to classify the test sample. If the test sample was unclassified, the RBC passed the associated antecedents onto the Statistic Based Classifier (SBC), if the SBC could not classify the test sample; the SBC passed the associated antecedents onto the General Inquirer Based Classifier (GIBC), which used the 3672 simple rules to determine the consequents of the antecedents. The Support vector machine (SVM) was given a training set to classify the test sample if the three classifiers failed to classify the same.

An ensemble technique is one which combines the outputs of several base classification models to form an integrated output. Rui Xia (2011) used this approach and made a comparative study of the effectiveness of ensemble technique for sentiment classification by efficiently integrating different feature sets and classification algorithms to synthesize a more accurate classification procedure. In his work, two types of feature sets are designed for sentiment classification, namely the part-of-speech based feature sets and the word-relation based feature sets. Then, three text classification algorithms, namely naive Bayes, maximum entropy and support vector machines, are employed as base-classifiers for each of the feature sets to predict classification scores. Three types of ensemble methods, namely the fixed combination, weighted combination and meta-classifier combination, are evaluated for three ensemble strategies namely ensemble of feature sets, ensemble of classification algorithms, and ensemble of both feature sets and classification algorithms.

In most of the comparative studies it is found that SVM outperforms other machine learning methods in sentiment classification. Ziqiong Zhang (2011) showed a contradiction in the performance of SVM. They focused their interest on written Cantonese, a written variety of Chinese. They proposed a method which utilizes completely prior-knowledge-free supervised machine learning method and proved that the chosen machine learning model could be able to draw its own conclusion from the distribution of lexical elements in a piece of Cantonese review. Despite its unrealistic independence assumption, the naive Bayes classifier surprisingly achieves better performance than SVM.

Sentiment classification is done by constructing a text classifier by extracting association rules that associate the terms of a document and its categories, by modeling the text documents as a collection of transactions where each transaction represents a text document, and the items in the transaction are the terms selected from the document and the categories the document is assigned to. Then, the system discovers associations between the words in documents and the labels assigned to them. Each category is considered as a separate text collection and the association rule mining is applied to it. The rules generated from all the categories separately are combined together to form the classifier (Weitong Huang, 2008). Then the training set is used to evaluate the classification quality, for classifying the test text documents, the number of rules covered, and attributive probability are used. Yulan He (2010) attempted to create a novel framework for sentiment classifier learning from unlabeled documents. The process begins with a collection of un-annotated text and a sentiment lexicon. An initial classifier is trained by incorporating prior information from the sentiment lexicon which consists of a list of words marked with their respective polarity. The labeled features use them directly to constrain model's predictions on unlabeled instances using generalized expectation criteria. The initially-trained classifier using generalized expectation is then applied on the un-annotated text and the documents labeled with high confidence are fed into the self-learned features extractor to acquire domain-dependent features automatically. Such self-learned features are subsequently used to train another classifier which is then applied on the test set to obtain the final results.

A few recent studies in this field explained the use of neural networks in sentiment classification. Zhu Jian (2010) proposed an individual model based on Artificial neural networks to divide the movie review corpus into positive , negative and fuzzy tone which is based on the advanced recursive least squares back propagation training algorithm. Long-Sheng Chen (2011) proposed a neural network based approach, which combines the advantages of the machine learning techniques and the information retrieval techniques.

3.2.Semantic Orientation

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

Much of the research in unsupervised sentiment classification makes use of lexical resources available. Kamps et al (2004) focused on the use of lexical relations in sentiment classification. Andrea Esuli and Fabrizio Sebastiani (2005) proposed semi-supervised learning method started from expanding an initial seed set using WordNet. Their basic assumption is terms with similar orientation tend to have similar glosses. They determined the expanded seed term's semantic orientation through gloss classification by statistical technique.

When the review where an opinion lies in, cannot provide enough contextual information to determine the orientation of opinion, Chunxu Wu(2009) proposed an approach which resort to other reviews discussing the same topic to mine useful contextual information, then use semantic similarity measures to judge the orientation of opinion. They attempted to tackle this problem by getting the orientation of context independent opinions , then consider the context dependent opinions using linguistic rules to infer orientation of context distinct-dependent opinion ,then extract contextual information from other reviews that comment on the same product feature to judge the context indistinct-dependent opinions.

An unsupervised learning algorithm by extracting the sentiment phrases of each review by rules of part-of-speech (POS) patterns was investigated by Ting-Chun Peng and Chia-Chun Shih (2010). For each unknown sentiment phrase, they used it as a query term to get top-N relevant snippets from a search engine respectively. Next, by using a gathered sentiment lexicon, predictive sentiments of unknown sentiment phrases are computed based on the sentiments of nearby known sentiment words inside the snippets. They consider only opinionated sentences containing at least one detected sentiment phrase for opinion extraction. Using the POS pattern opinion extraction is done. Gang Li & Fei Liu (2010) developed an approach based on the k-means clustering algorithm. The technique of TF-IDF (term frequency – inverse document frequency) weighting is applied on the raw data. Then, a voting mechanism is used to extract a more stable clustering result. The result is obtained based on multiple implementations of the clustering process. Finally, the term score is used to further enhance the clustering result. Documents are clustered into positive group and negative group.

Chaovalit and Zhou (2005) compared the Semantic Orientation approach with the N-gram model machine learning approach by applying to movie reviews. They confirmed from the results that the machine learning approach is more accurate but requires a significant amount of time to train the model. In comparison, the semantic orientation approach is slightly less accurate but is more efficient to use in real-time applications. The performance of semantic orientation also relies on the performance of the underlying POS tagger.

3.3.Role of negation

Negation is a very common linguistic construction that affects polarity and therefore, needs to be taken into consideration in sentiment analysis. Negation is not only conveyed by common negation words (not, neither, nor) but also by other lexical units. Research in the field has shown that there are many other words that invert the polarity of an opinion expressed, such as valence shifters, connectives or modals. “I find the functionality of the new mobile less practical”, is an example for valence shifter, “Perhaps it is a great phone, but I fail to see why”, shows the effect of connectives. An example sentence using modal is, “In theory, the phone should have worked even under water”. As can be seen from these examples, negation is a difficult yet important aspect of sentiment analysis.

Kennedy and Inkpen (2005) evaluate a negation model which is fairly identical to the one proposed by Polanyi and Zaenen (2004) in document-level polarity classification. A simple scope for negation is chosen. A polar expression is thought to be negated if the negation word immediately precedes it. Wilson et al. (2005) carry out more advanced negation modeling on expression-level polarity classification. The work uses supervised machine learning where negation modeling is mostly encoded as features using polar expressions. Jin-Cheon Na (2005), reported a study in automatically classifying documents as expressing positive or negative.He investigated the use of simple linguistic processing to address the problems of negation phrase.

In sentiment analysis, the most prominent work examining the impact of different scope models for negation is Jia et al. (2009). They proposed a scope detection method to handle negation using static delimiters, dynamic delimiters, and heuristic rules focused on polar expressions Static delimiters are unambiguous words, such as because or unless marking the beginning of another clause. Dynamic delimiters are, however, rules, using contextual information such as their pertaining part-of-speech tag. These delimiters suitably account for various complex sentence types so that only the clause containing the negation is considered. The heuristic rules focus on cases in which polar expressions in specific syntactic configurations are directly preceded by negation words which results in the polar expression becoming a delimiter itself.

3.4.Feature based sentiment classification

Due to the increasing amount of opinions and reviews on the internet, Sentiment analysis has become a hot topic in data mining, in which extracting opinion features is a key step. Sentiment analysis at both the document level and sentence level has been too coarse to determine precisely what users like or dislike. In order to address this problem, sentiment analysis at the attribute level is aimed at extracting opinions on products' specific attributes from reviews.

Hu’s work in (Hu, 2005) can be considered as the pioneer work on feature-based opinion summarization. Their feature extraction algorithm is based on heuristics that depend on feature terms’ respective occurrence counts. They use association rule mining based on the Apriori algorithm to extract frequent itemsets as explicit product features. Popescu et al (2005) developed an unsupervised information extraction system called OPINE, which extracted product features and opinions from reviews. OPINE first extracts noun phrases from reviews and retains those with frequency greater than an experimentally set threshold and then assesses those by OPINE’s feature assessor for extracting explicit features. The assessor evaluates a noun phrase by computing a Point-wise Mutual Information score between the phrase and meronymy discriminators associated with the product class. Popescu et al apply manual extraction rules in order to find the opinion words.

Kunpeng Zhang (2009), proposed a work which used a keyword matching strategy to identify and tag product features in sentences. Bing xu (2010) , presented a Conditional Random Fields model based Chinese product features identification approach, integrating the chunk features and heuristic position information in addition to the word features, part-of-speech features and context features.

Khairullah Khan et al (2010) developed a method to find features of product from user review in an efficient way from text through auxiliary verbs (AV) {is, was, are, were, has, have, had}. From the results of the experiments, they found that 82% of features and 85% of opinion-oriented sentences include AVs. Most of existing methods utilize a rule-based mechanism or statistics to extract opinion features, but they ignore the structure characteristics of reviews. The performance has hence not been promising.

Yongyong Zhail (2010) proposed a approach of Opinion Feature Extraction based on Sentiment Patterns, which takes into account the structure characteristics of reviews for higher values of precision and recall. With a self constructed database of sentiment patterns, sentiment pattern matches each review sentence to obtain its features, and then filters redundant features regarding relevance of the domain, statistics and semantic similarity.

Gamgarn Somprasertsri (2010) dedicated their work to properly identify the semantic relationships between product features and opinions. His approach is to mine product feature and opinion based on the consideration of syntactic information and semantic information by applying dependency relations and ontological knowledge with probabilistic based model.

4. Applications and Tools

Some of the applications of sentiment analysis includes online advertising, hotspot detection in forums etc.

Online advertising has become one of the major revenue sources of today’s Internet ecosystem. Sentiment analysis find its recent application in Dissatisfaction oriented online advertising Guang Qiu(2010) and Blogger-Centric Contextual Advertising (Teng-Kai Fan, Chia-Hui Chang ,2011), which refers to the assignment of personal ads to any blog page, chosen in according to bloggers’ interests.

When faced with tremendous amounts of online information from various online forums, information seekers usually find it very difficult to yield accurate information that is useful to them. This has motivated the research on identification of online forum hotspots, where useful information is quickly exposed to those seekers. Nan Li (2010) used sentiment analysis approach to provide a comprehensive and timely description of the interacting structural natural groupings of various forums, which will dynamically enable efficient detection of hotspot forums.

In order to identify potential risks, it is important for companies to collect and analyze information about their competitors' products and plans. Sentiment analysis find a major role in competitive intelligence (Kaiquan Xu , 2011) to extract and visualize comparative relations between products from customer reviews, with the interdependencies among relations taken into consideration, to help enterprises discover potential risks and further design new products and marketing strategies.

Opinion summarization summarizes opinions of articles by telling sentiment polarities, degree and the correlated events. With opinion summarization, a customer can easily see how the existing customers feel about a product, and the product manufacturer can get the reason why different stands people like it or what they complain about. Ku, Liang, and Chen (2006) investigated both news and web blog articles. Algorithms for opinion extraction at word, sentence and document level are proposed. The issue of relevant sentence selection is discussed, and then topical and opinionated information are summarized. Opinion summarizations are visualized by representative sentences. Finally, an opinionated curve showing supportive and non-supportive degree along the timeline is illustrated by an opinion tracking system.

Other applications includes online message sentiment filtering-mail sentiment classification, web blog author's attitude analysis etc.

Review Seer is a tool that automates the work done by aggregation sites. Naive Bayes classifier is used with positive and negative review sets for assigning a score to the extracted feature terms. The classifier did not perform well for web pages crawled from the result of a search engine. It displays attributes and score of the attribute along with review sentences.

Web Fountain uses beginning definite Base Noun Phrase (bBNP) heuristic for extracting product features. To assign sentiments to the features, reviews are parsed and traversed with two linguistic resources namely the sentiment lexicon and the sentiment pattern database. The sentiment lexicon defines the polarity of terms and sentiment pattern database defines sentiment extraction patterns for a sentence predicates (Yi and Niblack, 2005).

Red Opal is a tool that enables users to find products based on features. It scores each product based on features from the customer reviews (Christopher Scaffidi, 2007). Opinion observer is a sentiment analysis system for analyzing and comparing opinions on the web. The system shows the results in a graph format showing opinion of the product feature by feature (Bing Liu, 2005).

Besides these automated tools, various online tools like Twitrratr, Twendz ,Social mention, and Sentimetrics are available to track the sentiment in social networks.

5. Evaluation & Discussion

The performance of different methods used for opinion mining is evaluated by calculating various metrics like precision, recall and F-measure. Precision is the fraction of retrieved instances that are relevant, while recall is the fraction of relevant instances that are retrieved. The two measures are sometimes used together in the F_1 score (also F-score or F-measure) is a measure of a test's accuracy. An overview of the work done in the task of Sentiment Analysis is shown in Table 1. This table represents a sample of work done and some works published on the topic of Sentiment Analysis. It is evident from the Table 1,as far as the data source is concerned, a lot of work has been done on movie and product reviews. Internet Movie Database (IMDb) is used for movie reviews and product reviews are downloaded from Amazon.com.

Movie review mining is a more challenging application than many other types of review mining. The challenges of movie review mining lie in that factual information is always mixed with real-life review data and ironic words are used in writing movie reviews. Product review domain considerably differs from movie review domain because of two reasons. Firstly, there are feature specific comments in product reviews. People may like some features and dislike some others. Thus reviews consist of both positive and negative opinions, which make the task of classifying the review as positive or negative tougher. Such feature specific comments occur less frequently in movie reviews. Secondly, there are a lot of comparative sentences in product reviews and people talk about other products in reviews. This makes the task of opinion target detection an important aspect of the problem. A comparative analysis is done for sentiment analysis using movie review dataset (Fig 1) and product review from amazon.com (Fig 2) as data source.

Various methods have been used to measure the performance. From the performance achieved by these methods it is difficult to judge the best choice of classification method, since each method uses a variety of resources for training and different collections of documents for testing, various feature selection methods and different text granularity.

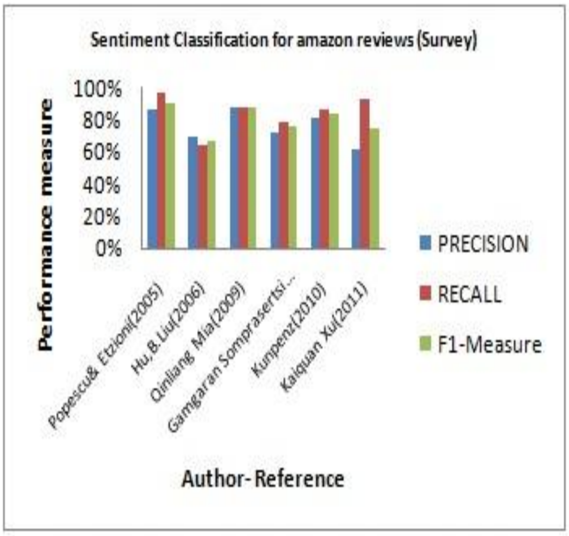

Fig 1. Sentiment Classification for Amazon reviews (Survey)

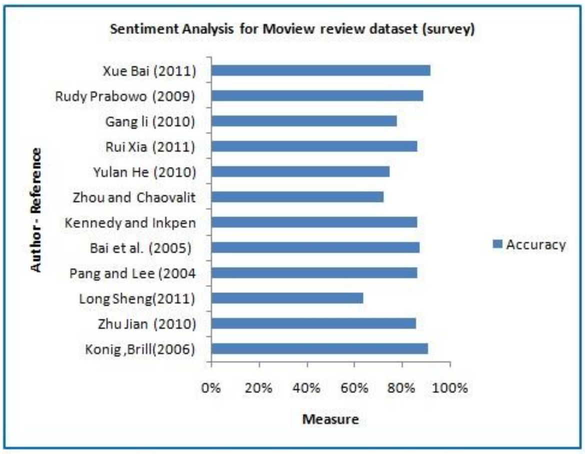

Fig 2. Sentiment Analysis for Moviereview dataset (survey)

6. Conclusion

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

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

Table 1. Summary of the survey

S.No	Studies	Mining Technique Used	Feature Selection	Data Source	Performance (Accuracy)	Precision	Recall	F1
1.	Kaiquan Xu(2011)	Multiclass SVM	Linguistic Feature	Amazon reviews	61%	61.9%	93.4%	74.2%
2.	Long Sheng(2011)	BPN	Point wise Mutual information	Movie review	64%	60%	98%	75%
3.	Rui Xia (2011)	Naïve bayes, Maximum entropy, SVM	Uni gram, bi grams, deendency grammar, joint feature	Movie review, Multi domain dataset,Amazon.	NB-85.8% ME-85.4% SVM-86.4%	-	-	-
4.	Xue Bai (2011)	Naïve bayes,	Information gain, two stage markov blanket classifier	Moview review,	92%	-	-	-
5.	Ziqiong (2011)	Naïve bayes,SVM	Information gain	Cantonese reviews	93%	-	-	-
6.	Gamgarn somprasti (2010)	Maximum Entropy	Dependency relation	Amazon reviews	-	72.6%	78.7%	75.4%
7.	Gang li (2010)	K-means Clustering	TF-IDF	Moview review	78%	-	-	-
8.	Yulan He (2010)	Sentiment lexicon, General expectation criteria	Self trained features	Movie review	74.7%	-	-	-
9.	Zhu Jian (2010)	Back propogation	Odds ratio	Movie review	86%	-	-	-
10.	Melville (2009)	Bayesian classification	n-grama	Blogs	Blogs: 91.21%	-	-	-
11.	Rudy(2009)	ID3,SVM,Hybrid	Document frequency	Moview review, mySpace comments	89%	-	-	-

Table 1. Summary of the survey (Continued)

S.No	Studies	Mining Technique Used	Feature Selection	Data Source	Performance (Accuracy)	Precision	Recall	F1
1.	QingliangMiao (2009)	Lexical resource	POS,Apriori	Amazon reviews	87.6%	87.4%	87.6%	87.4%
2.	Songho tan (2008)	Centroid classifier, K-Nearest neighbourhood,Winnow Classifier, SVM	MI,IG,CHI,DI	Chnsenticorp	90% (SVM)	-	-	-
3.	Zhou and Chaovalit (2008)	ontology-supported polarity mining	n-grams	Movie review	72.2%	-	-	-
4.	Godbole et al. (2007)	Lexical approach	Graph distance measurement	blog posts	82.7–95.7%	-	-	-
5.	Kennedy and Inkpen (2006)	support vector machines, termcounting	term frequencies	Movie review	86.2%	-	-	-
6.	Konig ,Brill(2006)	Hybrid	n-gram	Movie revie	91%	-	-	-
7.	Gamon (2005)	Naïve Bayes	Stemmed terms	Car reviews	86%	-	-	-
8.	Hu and Liu (2005)	Opinion word extraction and aggregation enhanced with	Opinion words	Amazon Cnn.Net	DVD- 73% MP3-93%	-	-	-
9.	Bai . (2005)	two-stage Markov Blanket Classier	dependence among words, minimal vocabulary	Movie review	87.5%	-	-	-
10.	Popescu and Etzioni (2005)	relaxation labeling clustering	templates, conjunctions and disjunctions,	Amazon	-	86%	97%	-
11.	Pang and Lee (2004	Nave Bayes support vector machines	based on minimum cuts	Movie review	86.4%	-	-	-

References

[1] Andrea Esuli and Fabrizio Sebastiani, “Determining the semantic orientation of terms through gloss classification”,Proceedings of 14th ACM International Conference on Information and Knowledge Management,pp. 617-624, Bremen, Germany, 2005.

[2] Bai, and R. Padman, “Markov blankets and meta-heuristic search: Sentiment extraction from unstructured text,” Lecture Notes in Computer Science, vol. 3932, pp. 167–187, 2006.

[3] Bing xu, tie-jun zhao, de-quan zheng, shan-yu wang, “Product features mining based on conditional random fields model “ , Proceedings of the Ninth International Conference on Machine Learning and Cybernetics, Qingdao, 11-14 July 2010

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

[5] Chau, M., & Xu, J. (2007). Mining communities and their relationships in blogs: A study of online hate groups. International Journal of Human – Computer Studies, 65(1), 57–70.

[6] Christopher Scaffidi, Kevin Bierhoff, Eric Chang, Mikhael Felker, Herman Ng and Chun Jin, “Red Opal: product-feature scoring from reviews”, Proceedings of 8th ACM Conference on Electronic Commerce, pp. 182-191, New York, 2007.

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

[8] Gamgarn Somprasertsri, Pattarachai Lalitrojwong , Mining Feature-Opinion in Online Customer Reviews for Opinion Summarization, Journal of Universal Computer Science, vol. 16, no. 6 (2010), 938-955.

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

[10] Go,Lei Huang and Richa Bhayani , “Twitter Sentiment Analysis”, Project Report, standford,2009.

[11] Go,Lei Huang and Richa Bhayani , “Twitter Sentiment Classification using Distant Supervision”, Project Report, Standford,2009.

[12] Guang Qiu , Xiaofei He , Feng Zhang , Yuan Shi , Jiajun Bu , Chun Chen , “DASA: Dissatisfaction-oriented

Advertising based on Sentiment analysis” , Expert Systems with Applications, 37 (2010) 6182–6191.

[13] Hu, and Liu, “Opinion extraction and summarization on the web”, AAAI., (2006), pp. 1621-1624.

[14] Hu, and Liu, “Mining and summarizing customer reviews”, Proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, WA, USA, 2005,pp. 168–177.

[15] Hu, Liu and Junsheng Cheng, “Opinionobserver: analyzing and comparing opinions on theWeb”, Proceedings of 14th international Conference onWorldWideWeb, pp. 342-351, Chiba, Japan, 2005.

[16] Jia, C. Yu, andW. Meng , “The Effect of Negation on Sentiment Analysis and Retrieval Effectiveness”, In Proceedings of CIKM ,2009.

[17] Jin-Cheon Na , Christopher Khoo, Paul Horng Jyh Wu, “Use of negation phrases in automatic sentiment classification of product reviews”, Library Collections, Acquisitions,& Technical Services 29 (2005) 180–191.

[18] Kaiquan Xu , Stephen Shaoyi Liao , Jiexun Li, Yuxia Song, “Mining comparative opinions from customer reviews for Competitive Intelligence”, Decision Support Systems 50 (2011) 743–754.

[19] Kamps, Maarten Marx, Robert J. Mokken and Maarten De Rijke, “Using wordnet to measure semantic orientation of adjectives”, Proceedings of 4th International Conference on Language Resources and Evaluation, pp. 1115-1118, Lisbon, Portugal, 2004.

[20] Kennedy and D. Inkpen, “Sentiment classification of movie reviews using contextual valence shifters,” Computational Intelligence, vol. 22, pp. 110–125,2006.

[21] Khairullah Khan, Baharum B. Baharudin, Aurangzeb Khan, and Fazal_e_Malik, “Automatic Extraction of Features and Opinion Oriented Sentences from Customer Reviews”, World Academy of Science, Engineering and Technology 62 2010.

[22] Ku, L.-W., Liang, Y.-T., & Chen, H, “Opinion extraction, summarization and tracking in news and blog corpora”. In AAAI-CAAW’06.

[23] Kunpeng Zhang, Ramanathan Narayanan, “Voice of the Customers: Mining Online Customer Reviews for Product”, 2010.

[24] Long-Sheng Chen , Cheng-Hsiang Liu, Hui-Ju Chiu , “A neural network based approach for sentiment classification in the blogosphere”, Journal of Informetrics 5 (2011) 313–322.

[25] Martin, J. (2005). Blogging for dollars. Fortune Small Business, 15(10), 88–92.

[26] Melville, Wojciech Gryc, “Sentiment Analysis of Blogs by Combining Lexical Knowledge with Text Classification”, KDD’09, June 28–July 1, 2009, Paris, France.Copyright 2009 ACM 978-1-60558-495-9/09/06.

[27] Namrata Godbole, Manjunath Srinivasaiah, Steven Skiena, “LargeScale Sentiment Analysis for News and Blogs”, ICWSM’2007 Boulder, Colorado, USA.

[28] Nan Li , Desheng Dash Wu , “Using text mining and sentiment analysis for online forums hotspot detection and forecast”, Decision Support Systems 48 (2010) 354–368.

[29] Polanyi and A. Zaenen, “Contextual lexical valence shifters,” in Proceedings of the AAAI Spring Symposium on Exploring Attitude and Affect in Text,AAAI technical report SS-04-07, 2004.

[30] Popescu, A. M., Etzioni, O.: Extracting Product Features and Opinions from Reviews, In Proc. Conf. Human Language Technology and Empirical Methods in Natural Language Processing, Vancouver, British Columbia, 2005, 339–346.

[31] Qiang Ye, Ziqiong Zhang, Rob Law, “Sentiment classification of online reviews to travel destinations by supervised machine learning approaches”, Expert Systems with Applications 36 (2009) 6527–6535.

[32] Qingliang Miao, Qiudan Li, Ruwei Dai , “AMAZING: A sentiment mining and retrieval system”, Expert Systems with Applications 36 (2009) 7192–7198.

[33] Rudy Prabowo, Mike Thelwall, “Sentiment analysis: A combined approach .”, Journal of Informetrics 3 (2009) 143–157.

[34] Rui Xia , Chengqing Zong, Shoushan Li, “Ensemble of feature sets and classification algorithms for sentiment classification”, Information Sciences 181 (2011) 1138–1152.

[35] Songbo Tan , Jin Zhang, “An empirical study of sentiment analysis for chinese documents ”, Expert Systems with Applications 34 (2008) 2622–2629.

[36] Teng-Kai Fan, Chia-Hui Chang , “Blogger-Centric Contextual Advertising” Expert Systems with Applications 38 (2011) 1777–1788 .

[37] Ting-Chun Peng and Chia-Chun Shih , “An Unsupervised Snippet-based Sentiment Classification Method for Chinese Unknown Phrases without using Reference Word Pairs”, 2010 IEEE/WIC/ACM International Conference on Web Intelligence and ntelligent Agent Technology JOURNAL

OF COMPUTING, VOLUME 2, ISSUE 8, AUGUST 2010, ISSN 2151-9617 .

[38] Tsur, D. Davidov, and A. Rappoport , " A Great Catchy Name: Semi-Supervised Recognition of Sarcastic Sentences in Online Product Reviews". In Proceeding of ICWSM. of Context Dependent Opinions,2010.

[39] Weitong Huang, Yu Zhao, Shiqiang Yang, Yuchang Lu, "Analysis of the user behavior and opinion classification based on the BBS" , Applied Mathematics and Computation 205 (2008) 668–676 .

[40] Wiebe, W.-H. Lin, T. Wilson and A. Hauptmann, "Which side are you on? Identifying perspectives at the document and sentence levels," in Proceedings of the Conference on Natural Language Learning (CoNLL), 2006.

[41] Yi and Niblack, "Sentiment Mining in Web Fountain",Proceedings of 21st international Conference on Data Engineering, pp. 1073-1083, Washington DC,2005.

[42] Yongyong Zhail, Yanxiang Chenl, Xuegang Hu, "Extracting Opinion Features in Sentiment Patterns" , International Conference on Information, Networking and Automation (ICINA),2010.

[43] YuanbinWu, Qi Zhang, Xuanjing Huang, LideWu, "Phrase Dependency Parsing for Sentiment analysis", Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, pages 1533–1541, Singapore, 6-7 August 2009 .

[44] ZHU Jian , XU Chen, WANG Han-shi, "" Sentiment classification using the theory of ANNs", The Journal of China Universities of Posts and Telecommunications, July 2010, 17(Suppl.): 58–62 .

[45] Ziqiong Zhang, Qiang Ye, Zili Zhang, Yijun Li, "Sentiment classification of Internet restaurant reviews written in Cantonese", Expert Systems with Applications xxx (2011) xxx–xxx.