Opinion Mining and Sentiment Analysis –An Assessment of Peoples' Belief: A Survey

S Padmaja[1] and Prof. S Sameen Fatima[2]

[1]Research Scholar, Department of CSE, OUCE, Osmania University, Hyderabad, India.
bandupadmaja@gmail.com

[2]Professor and Head, Department of CSE, OUCE, Osmania University, Hyderabad, India.
sameenf@gmail.com

Abstract

Opinion Mining is a process of automatic extraction of knowledge from the opinion of others about some particular topic or problem. The idea of Opinion mining and Sentiment Analysis tool is to "process a set of search results for a given item, generating a list of product attributes (quality, features etc.) and aggregating opinion". But with the passage of time more interesting applications and developments came into existence in this area and now its main goal is to make computer able to recognize and generate emotions like human. This paper will try to focus on the basic definitions of Opinion Mining, analysis of linguistic resources required for Opinion Mining, few machine learning techniques on the basis of their usage and importance for the analysis, evaluation of Sentiment classifications and its various applications.

Keywords- *Sentiment Mining, Opinion Mining, Text Classification.*

1. Introduction

Human life is filled with emotions and opinions. We cannot imagine the world without them. Emotions and opinions play a vital role in nearly all human actions. They lead the human life by influencing the way we think, what we do and how we act. Having an access to large quantities of data through internet and its transformation into a social web is no longer an issue, as there are terabytes of new information produced on the web everyday that are available to any individual. Even more importantly, it has changed the way we share information. The receivers of the information do not only consume the available content on web, but in turn, actively annotate this content and generate new pieces of information. Today people not only comment on the existing information, bookmark pages and provide ratings but they also share their ideas, news and knowledge with the community at large. In this way, the entire community becomes a writer, in addition to being a reader. The existing mediums like Blogs, Wikis, Forums and Social Networks where users can post information, give opinions and get feedback from other users on different topics, ranging from politics and health to product reviews and travelling. The increasing popularity of personal publishing services of different kinds suggests that opinionated information will become an important aspect of the textual data

on the web. Recently, many researchers have focused on this area. They are trying to fetch opinion information to analyze and summarize the opinions expressed automatically with computers. This new research domain is usually called Opinion Mining and Sentiment Analysis. Until now, researchers have evolved several techniques to the solution of the problem. Current-day Opinion Mining and Sentiment Analysis is a field of study at the crossroad of Information Retrieval (IR) and Natural Language Processing (NLP) and share some characteristics with other disciplines such as text mining and Information Extraction.

The rest of the paper is organized as follows. In section 2 the basic definitions have been discussed. Section 3 gives the overview of linguistic resources in Opinion Mining. Section 4 covers some important machine learning techniques which are commonly used in Sentiment Analysis. Section 5 presents various measures of evaluating Sentiment Classification. Section 6 showcases the wide range of its applications while section 7 highlights various NLP tools that are commonly used for Sentiment Analysis and finally section 8 concludes the paper.

2. Basic Definitions

2.1 Subjectivity Analysis: A general view

Subjectivity Analysis involves various methods and techniques that originate from IR, Artificial Intelligence and NLP. This confluence of different approaches is explained by the nature of the data being processed and application requirements. The subjectivity Analysis domain is still in the process of being shaped and its problem statements touch upon different domains mentioned above. Moreover, Opinion Mining originates from the IR community, and aims at extracting and further processing users' opinions about products, movies or other entities. Sentiment analysis, on the other hand, was initially formulated as the NLP task of retrieval of sentiments expressed in texts. However these two problems are similar in their own essence and fall under the scope of Subjectivity Analysis.

2.2. Definition of Opinion

An Opinion is a belief or judgment of a large number or majority of people formed about a particular thing, not necessarily based on fact or knowledge. In general, opinion refers to what a person thinks about something. In other words, opinion is a subjective belief, and is the result of emotion or interpretation of facts.

2.3. Document, Topic and Sentiment

A Document D is a piece of text in natural language. We assume that each document discusses at least one topic, and not all topics discussed in the same document have to be related to each other. Topic T is a named entity, event or abstract concept that is mentioned in a document D and a Sentiment S is the author's attitude, opinion or emotion expressed on topic T.

2.4. Opinion mining or Sentiment analysis

Opinion mining is a technique to detect and extract subjective information in text documents. In general, sentiment analysis tries to determine the sentiment of a writer about some aspect or the overall contextual polarity of a document. The sentiment may be his or her judgment, mood or evaluation. A key problem in this area is sentiment classification, where a document is labeled as a positive or negative evaluation of a target object (film, book, product etc.)

3. Analysis of linguistic resources for Opinion Mining

The basic problem of opinion mining is opinion extraction. It is required to know the linguistic terms and get the idea from the text classification of contents of document into positive and negative and subjective and objective terms identified by syntactic features. Another main focus is on subjectivity detection. Subjectivity is used to express private states in the context of a text or conversation. Private state is a general term for opinions, evaluation, beliefs, perception, emotions etc. Objective sentiment conveys information in accordance with the intension of the author. If a user feedback has no judgment or opinion on the source content then it is called objective. Changli Zhang et al. [27] in their work have used Bag-of-Word (BOW) and appraisal phrase and get 79.0% result through BOW and 80.26% with the combination of BOW and appraisal phrase [2]. In [28], Minqing Hu and Bing Liu have used Natural Language Processor linguistic parser to parse each review to split text into sentences and to produce part of speech tags for each word like noun, verb, adjective etc. Some authors have taken term senses into account and assume that a single term can be used in a different sense and can present a different opinion. WordNet and Synsets are used to examine different senses of the same term.

3.1 Text features identification and Orientation

The text features identification has three different levels. They are words, sentences and documents. Existing research work presents different techniques and ideas for extraction of sentimental terms from text. According to linguistic rules words and phrases are categorized as noun, verbs, adjectives and adverbs. Most of the work uses part of speech (POS), stop word removal, fuzzy pattern matching, stemming, punctuation, link based patterns, document citation and stylistic measures for extraction of sentiments.[16, 17, 18]

3.2 Adjectives, Nouns, Verbs and Adverbs

Existing research of polarity classification mainly focus on adjectives and adverbs to identify subjectivity [19, 20, 21]. From experiments they have shown that opinion extraction using adjective has precision of 64.2% and a recall of 69.3%. Most commonly used tool for adjective identification is WordNet [22]. Farah Benamara et al. Have proposed that adjective and adverbs are better than adjectives alone [23]. In most of the existing work, sentiment expressions mainly depend on some words which can express subjective sentiment orientation. For example good is used for positive and bad is used for negative sentiment orientation. Such subjective words are actually called adjectives in linguistic terms. Verb identification plays an important role in finding relationship between subjective and objective terms. According to Turney, Adjectives, Nouns, Verbs and Adverbs are grammatical categories which have the capacity to express emotions and subjectivity   [24].

3.3 Semantic Orientation of Text

Classification of sentimental expression according to their meaning and background knowledge is called Semantic Orientation. Though Syntactic analysis plays a key role in document classification, it is not sufficient to extract the concept from the text only through syntax. T Hoffmann combined information theoretic measures and semantic knowledge of a hierarchy using WordNet to extract concept from text automatically. Turney [24] and Pu Wang et al [25] have used Bag of Word (BOW) and semantic

concepts to enrich the representation of text classification and to extract concept from text.

3.4 Ontology Based Learning

Ontology can be defined as a formal knowledge representation system (KRS) which has three main components. Classes (or concepts or topics), instances which are individual which belong to a class) and properties (which link classes and instances allowing to insert information regarding the world into the ontology).Ontology based learning is a growing area of research for extracting opinion from text. It integrates the domain knowledge of individual words into the terms for learning and capturing concept from text. The relationship between terms in text is helpful in understanding the background knowledge. Wen Zhang et al [26] have worked on text classification based on multi word using ontology.

4. Commonly used Machine Learning Models for text Classification

4.1 Naïve Bayesian Classifications

Naïve Bayesian method is one of the popular techniques for text classification. It has been shown to perform extremely well in practice by many researchers [29, 30]. Given a set of training documents

D, each document is considered an ordered list of words. Let $w_{di, k}$ denotes the word in position k of document di, where each word is from the vocabulary V= $<w_1,w_2…,w_{|v|}>$,where vocabulary is the set of all words we consider for classification, and let a set of pre-defined classes be C= $<c_1, c_2…, c_{|C|}>$. In order to perform classification, we need to compute the posterior probability, $P[c_j|d_i]$. Based on the Bayesian probability and the multinomial model, we have

$$P[cj] = \sum_i P[cj\,/\,di]\,/D.$$

To eliminate zeros, Laplacian smoothing can be used [31], which simply adds one to each count:

$$P[wt\,/\,cj] = \frac{1 + \sum_{i-1}^{D} N(wt, di)P(cj\,/\,di)}{V + \sum_{s-1}^{V} \sum_{i-1}^{D} N(ws, di)P(cj\,/\,di)}$$

where $N(w_s, d_i)$ is the number of times the word, w_t, occurs in document, d_i, and $P(c_j|d_i)$ {0,1} depends on the class label of the document. Finally, assuming that the probabilities of the words are
Independent given the class, we obtain

$$P[cj\,/\,di] = \frac{P[cj]\prod_{k-1}^{di} P[wdi.k\,/\,cj]}{\sum_{r-1}^{C} P[cr]\prod_{k-1}^{di} P(wdi.k\,/\,cr)}$$

In the naïve Bayes classifier, the class with the highest P [cj|di] is assigned as the class of the document. Thus it is a supervised learning method. A Bayesian Classifier is a simplest

probabilistic classifier based on Bayes theorem. In text classification, to determine the most probable class or group, a document falls into, Bayes rule is used.

4.2 Support Vector Machine (SVM)

As explained in Dumais and Chen (2000) and Pang et al (2002). Given a category set, C = {+1, −1} and two pre-classified training sets, i.e., a positive sample set, $Tr^+ = \sum_{i=1}^{n}$ (d_i, +1)

and a negative sample set, $Tr^- = \sum_{i=1}^{n}$ (d_i, −1), the SVM finds a hyperplane that separates the two sets with maximum margin (or the largest possible distance from both sets), as illustrated in Fig. 1. At pre-processing step, each training sample is converted into a real vector, xi that consists of a set of significant features representing the associated document, di. Hence, Tr + $= \sum_{i=1}^{n}$ (x_i, +1) for the positive sample set and Tr − $= \sum_{i=1}^{n}$ (x_i, −1) for the negative sample set. In this regard, for c_i =+1, $w \cdot x_i + b > 0$, and for c_i = -1, $w \cdot x_i + b < 0$. Hence, T +, T − {$c_i \cdot (w \cdot x_i + b) \geq 1$} becomes an optimization problem defined as follows:

minimize (1/2) ||w||2 , subject to $c_i \cdot (w \cdot x_i + b) \geq 1$. The result is a hyperplane that has the largest distance to x_i from both sides. The classification task can then be formulated as discovering which side of the hyperplane a test sample falls into [15].
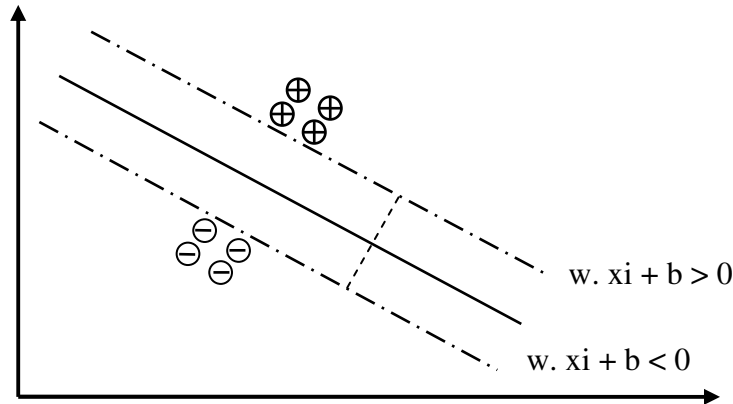

w. xi + b > 0

w. xi + b < 0

Fig 1: An illustration of the SVM method

4.3 K-nearest neighbor (KNN)

KNN is a simple machine learning algorithm. In this algorithm, the objects are classified based on the majority of its neighbor. The class assigned to the object is most among its k nearest neighbors. The KNN classification algorithm classifies the instances or objects based on their similarities to instances in the training data. In KNN, selection is based on majority voting or distance weighted voting. KNN is unsupervised text classification algorithm and it works efficiently when the training set is large. Consider the vector A and set of M labeled instances {ai, bi}. The classifier predicts the class label of A on the predefined N classes. The KNN classification algorithm finds the k nearest neighbors of A and determines the class label of A using majority vote. KNN classifier applies Euclidean distances as the distance metric [1].

$$\text{Dist}(X, Y) = \sqrt{\sum (X_i - Y_i)^2}$$

4.4 Maximum entropy classification

In Maximum Entropy classification, the probability that a document belongs to a particular class given a context must maximize the entropy of the classification system. By maximizing entropy, it is ensured that no biases are introduced into the system. The model makes no assumptions of the independence of words. However, it is computationally more expensive. It is a machine learning method based on empirical data. Nigam et al [32] and Berger et al [33] showed that in many cases it outperforms Naïve Baye's classification. Raychaudhari et al [34] also found that Maximum Entropy worked better than Naïve Baye's and Nearest Neighbor classification for their classification. Unlike the Naïve Baye's machine learning, Maximum Entropy makes no independent assumptions about the occurrence of words. The Maximum Entropy modeling technique provides a probability distribution that is as close to the uniform as possible given that the distribution satisfies certain constraints. We provide only a terse overview of Maximum entropy. A full description of the method can be found in Manning and Schutze [35] and Ratnaparkhi [36].

The classification system is well described by Ratnaparkhi [36] as: "Maximum Entropy models offer a way to combine diverse pieces of contextual evidence in order to estimate the probability of a certain linguistic class occurring with a certain linguistic context….in which task is to estimate the probability of class 'a' occurring with context 'b' ".

The principle of the Maximum Entropy modeling states that: "The Maximum Entropy probability distribution, P*, is the unique distribution that maximizes:

$$H = \Sigma P^*(V) \log (P^*(V)) \ \forall V$$

While satisfying the supplied constraints" [37]. The Maximum Entropy classification requires a set of features, which define a category. For example, in case of documents, features could be the words that belong to the documents in that category. A feature f is a binary function that maps to '1' if a document belonging to a category contains the feature (word). Thus:

$$f = 1 \text{ iff } \text{"ABC"} \in d \text{ and } c = \text{"XYZ"}$$

The probability that a document belongs to a particular category is given by:

$$P(cj/d) = \frac{1}{Z(d)} \exp\left(\sum_i \lambda i f i(d, cj) \right)$$

Where $P(c_j|d)$ is the probability that a class occurs for a given document. $Z(d)$ is the normalizing constant that is obtained by summing over all $P(c_j \mid d)$ over all values of j. The probability distribution $P^*(V)$ is calculated by an iterative method called Generalized Iterative Scaling, which begins with a representation of the uniform distribution and converges towards the maximum entropy distribution. The values of λ_i are obtained so that the system satisfies the constraint that the observed expectation of a feature in the universe should match the expectation of the feature in the given sample set.

The motivation behind maximum entropy is that, for certain data, one should prefer the most uniform models that also satisfy any given constraint. The main advantage of maximum entropy is being able to combine multiple knowledge sources and adding additional knowledge easily. In general formulation, maximum entropy can be used to estimate any probability distribution. It is an optimization problem [4].

4.5 Boosting Algorithm

Boosting (Schapire, 1990) is a meta-algorithm which can be viewed as a model averaging method. It is the most widely used ensemble method and one of the most powerful learning ideas introduced in the last ten years. Originally designed for classification, it can also be profitably extended to regression. A "weak" classifier is created, that is, it suffices that its accuracy on the training set is only slightly better than a random guessing. A succession of models are built iteratively, each one being trained on a data set in which points misclassified (or, with regression, those poorly predicted) by the previous model are given more weight. Finally, all of the successive models are weighted according to their success and then the outputs are combined using voting (for classification) or averaging (for regression), thus creating a final model.

Table 1: An overview of the most popular machine learning algorithms used in Sentiment Analysis

ML Algorithm	Type	Study	Year	Paper	Dataset	Accuracy (%)
		Annett and Kondrak	2008	[1]	Reviews	77.5
		Bifet and Frank	2010	[3]	Microblogs	82.5
		Chen et al.	2006	[6]	Reviews	77.5
		Dave et al.	2003	[7]	Reviews	81.9-87.0
Naïve Bayes	**Supervised**	gindl and Liegl	2008	[9]	Reviews	66.0
		Go et al.	2009	[10]	Microblogs	82.7
		Pang et al.	2002	[12]	Reviews	81.5
		Zhang et al.	2011	[14]	Reviews	84.5
		Annett and Kondrak	2008	[1]	Reviews	77.4
Support Vector Machines	**Supervised**	Chen et al.	2006	[6]	Reviews	84.6
		Go et al.	2009	[10]	Microblogs	82.2
		Pang et al.	2002	[12]	Reviews	82.9
		Shimada and Endo	2008	[13]	Reviews	77.1
Maximum entropy	**Supervised**	Gindl and Liegl	2008	[9]	Reviews	83.8
		Go et al.	2009	[10]	Microblogs	83
		Pang et al.	2002	[12]	Reviews	81.0
Boosting	**Supervised**	Kudo and Matsumoto	2004	[11]	Reviews	59.6-90.2
		Cassinelli and Chen	2009	[5]	Reviews	73.0-76.0
K-nearest neighbor	**UnSupervised**	Davidov et al.	2010	[8]	Microblogs	66.0-87.0

5. Evaluation of Sentiment Classification

In general, the performance of sentiment classification is evaluated by using four indexes. They are Accuracy, Precision, Recall and F1-score. The common way for computing these indexes is based on the confusion matrix as shown below:

Table 2: Confusion Matrix

#	Predicted positives	Predicted negatives
Actual positive instances	Number of True Positive instances (TP)	Number of False Negative instances (FN)
Actual negative instances	Number of False Positive instances (FP)	Number of True Negative instances (TN)

These indexes can be defined by the following equations:

- $\text{Accuracy} = \dfrac{TN + TP}{TN + TP + FP + FN}$

- $\text{Precision} = \dfrac{TP}{TP + FP}$

- $\text{Recall} = \dfrac{TP}{TP + FN}$

- $\text{F1} = \dfrac{2 \times \Pr ecision \times \mathrm{Re} call}{\Pr ecision + \mathrm{Re} call}$

Accuracy is the portion of all true predicted instances against all predicted instances. An accuracy of 100% means that the predicted instances are exactly the same as the actual instances. Precision is the portion of true positive predicted instances against all positive predicted instances. Recall is the portion of true positive predicted instances against all actual positive instances. F1 is a harmonic average of precision and recall.

6. Applications

Although the field of Sentiment Analysis is relatively young, there are already numerous businesses that use the techniques developed in this field to customers interested in brand tracking and market perception. Specifically, the various types of activities that may be involved are:

- Tracking collective user opinions are rating of products and services

- Analyzing consumer trends, competitors and market buzz
- Measuring response to company-related events and incidents
- Monitoring critical issues to prevent negative viral effects
- Evaluating feedback in multiple languages

As a source of opinionated discourse, these companies look at

- Online communities
- Discussion boards
- Weblogs
- Product rating sites
- Chat rooms
- Newsgroups

The above wide range of applications can briefly be categorized into information systems including review classification, review summarization, synonyms and antonyms extraction, opinions tracking in online discussions etc.

7. Commonly Used NLP Tools for Sentiment Analysis

A variety of open-source text-analytics tools like natural-language processing for information extraction and classification can be applied for sentiment analysis. The tools listed below can work on textual sources only.

- **LingPipe –** It is a suite of java tools for linguistic processing of text including entity extraction, speech tagging (pos), clustering, classification, etc. It is one of the most mature and widely used open source NLP toolkits in industry. It is known for its speed, stability, and scalability. One of its best features is the extensive collection of well-written tutorials which helps to get started. LingPipe is released under a royalty-free commercial license that includes the source code, but it's not technically 'open-source'. http://alias-i.com/lingpipe/demo/

- **OpenNLP** - hosts a variety of java-based NLP tools which perform the most common NLP tasks, such as tokenization, sentence segmentation, part-of-speech tagging, named entity extraction, chunking, parsing, and co-reference resolution. These tasks are usually required to build more advanced text processing services. OpenNLP also includes maximum entropy based machine learning. http://opennlp.apache.org/

- **Stanford Parser and Part-of-Speech (POS) Tagger** - Java packages for sentence parsing and part of speech tagging from the Stanford NLP group. It has implementations of probabilistic natural language parsers. http://nlp.stanford.edu/software/tagger.shtm/

- **NTLK** - The natural language toolkit is a tool for teaching and researching classification, clustering, speech tagging and parsing, and more. It contains a

set of tutorials and data sets for experimentation. It is written by Steven Bird, from the University of Melbourne. http://www.nltk.org/, http://text-processing.com/demo/...

- **Opinion Finder** – Opinion Finder, which was initially released in 2006, employs a multi-stage NLP process. It aims to identify subjective sentences and to mark various aspects of subjectivity in these sentences, including the source (holder) of the subjectivity and words that are included in phrases expressing positive or negative sentiments. http://code.google.com/p/opinionfinder/

- **Tawlk/osae** - A python library for sentiment classification on social text. The end-goal is to have a simple library that "just works". It should have an easy barrier to entry and be thoroughly documented. https://github.com/Tawlk/osae/

- **GATE** - GATE is over 15 years old and is in active use for all types of computational task involving human language. GATE excels at text analysis of all shapes and sizes. From large corporations to small startups, from multi-million research consortia to undergraduate projects. http://gate.ac.uk/sentiment/

- **textir** - A suite of tools for text and sentiment mining. This includes the 'mnlm' function, for sparse multinomial logistic regression, 'pls', a concise partial least square routine, and the 'topics' function, for efficient estimation and dimension selection in latent topic models. http://cran.r-project.org/web/packages/textir/index.html/

- **NLP Tool suite** – A comprehensive NLP tool suite is used for the application purposes of semantic search, information extraction and text mining. Most of their continuously expanding tool suite is based on machine learning methods and thus is domain and language independent. http://www.julielab.de/Resources/Software/NLP_Tools.html/

8. Conclusion

This paper introduced and surveyed the field of sentiment analysis and opinion mining. It tried to showcase from basic definitions, different techniques, various evaluation methods, wide range of applications to variety of NLP tools that are commonly used for Sentiment Analysis. It has been a very active research area in recent years. In fact, it has spread from computer science to management science. Finally, this paper concludes saying that all the sentiment analysis tasks are very challenging.

REFERENCES

[1] M.Annett, G. Kondrak, (2008), "A comparison of sentiment analysis techniques: polarizing movie Blogs", in: Advances in Artificial Intelligence, pg: 25-35.

[2] Khairullah Khan, Baharum B. Baharuddin, (2009), "Mining Opinion from Text Documents: A Survey", In: 3rd IEEE International Conference on Digital Ecosystem and Techniques.

[3] A. Bifet, E. Frank, (2010), "Sentiment knowledge discovery in twitter streaming data", in: Discovery Science, pg: 1-15.

[4] A. Nisha Jebaseel and Dr. E. Kirubakaran, (2012), "M-Learning Sentiment Analysis with Data Mining Techniques" in: International Journal of Computer Science and Telecommunications, Volume 3, Issue 8.

[5] A. Cassinelli, C.W. Chen, (2009), "Final Project Boost up! Sentiment Categorization with Machine Learning Techniques".

[6] C. Chen, F. Ibekwe-SanJuan, E. SanJuan, C. Weaver, (2006), "Visual analysis of conflicting opinions", in: Visual Analytics Science and Technology, IEEE Symposium On, pg. 59–66.

[7] K. Dave, S. Lawrence, D.M. Pennock, (2003), "Mining the peanut gallery: opinion extraction and semantic Classification of product reviews", in: Proceedings of the 12th International Conference on World Wide Web, pg. 519 – 528.

[8] D. Davidov, O. Tsur, A. Rappoport, (2003), "Enhanced sentiment learning using twitter hashtags and smileys", in: Proceedings of the 23rd International Conference on Computational Linguistics, Posters, pg. 241–249.

[9] S. Gindl, J. Liegl, (2008), "Evaluation of different sentiment detection methods for polarity classification on web-Based reviews", in: Computational Aspects of Effectual and Emotional Interaction.

[10] A. Go, R. Bhayani, L. Huang, (2009), "Twitter sentiment classification using distant supervision", in: CS224N Project Report, Stanford, pg. 1–12.

[11] T. Kudo, Y. Matsumoto, (2004), "A boosting algorithm for classification of semi-structured Text", in: Proceedings of EMNLP.

[12] B. Pang, L. Lee, S. Vaithyanathan, (2002), "Thumbs up? Sentiment classification using Machine learning Techniques", in: Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing, Volume 10, pg. 79–86.

[13] K. Shimada, T. Endo, (2002), "Seeing several stars: a rating inference task for a document containing several Evaluation criteria", in: Proceedings of the 12th Pacific-Asia Conference on Advances in Knowledge Discovery And Data Mining, pg. 1006–1014.

[14] Z. Zhang, Q. Ye, Z. Zhang, Y. Li, (2011), "Sentiment classification of Internet restaurant Reviews written in Cantonese, Expert Systems with Applications" pg: 7674–7682.

[15] Rudy Prabowo, Mike Thelwall, (2009), "Sentiment Analysis: A combined approach", Journal of Informatics pg: 143-157.

[16] Janyce Wiebe et al., (2003), "Recognizing and Organizing Opinions Expressed in the World Press", AAAI Spring Symposium on New Directions in Question Answering.

[17] Janyce Wiebe et al., (2002), "NRRC Summer Workshop on Multiple-Perspective Question Answering: Final Report", Theresa Wilson.

[18] David Pierce et al, (2010), "User-Oriented Machine Learning Strategies for Information Extraction: Putting the Human Back in the Loop", Working Notes of the IJCAI-2001 Workshop on Adaptive Text Extraction and Mining.

[19] Wiebe, J., (2000), "Learning subjective adjectives from corpora", Proceedings of the 17th National Conference on Artificial Intelligence (AAAI-2000), Austin, Texas.

[20] P. Chesley, B. Vincent, L. Xu, and R. Srihari, (2006), "Using verbs and adjectives automatically classify blogs sentiment," in AAAI Symposium on Computational Approaches to Analyzing Weblogs (AAAI-CAAW), pg. 27–29.

[21] V. Hatzivassiloglou and K. McKeown, (1997), "Predicting the semantic orientation of adjectives," in Proceedings of the Joint ACL/EACL Conference, pg. 174–181.

[22] C. Fellbaum, ed., (1998), "Wordnet: An Electronic Lexical Database", MIT Press.

[23] Farah Benamara, Carmine Cesarano, Diego Reforgiato, (2007), "Sentiment Analysis: Adjectives and Adverbs are better than Adjectives Alone", ICWSM ' Boulder, CO USA.

[24] Turney, (2002), "Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews", Proceedings of the 40th Annual Meeting of the Association of Computational Linguistics (ACL'02), pg: 417-424.

[25] Turney P. and Littman, (2003), "Measuring praise and criticism: Inference of semantic Orientation from association", ACM Transactions on Information Systems, Volume 21, Issue 4, pg: 315-346.

[26] Wen zhang , Taketoshida, and Xijin Tang, (2008), "Text classification based on multi-word with support vector machine" Elsevier knowledge-based systems.

[27] Changli Zhang et al., (2008), "Sentiment Classification for Chinese Reviews Using Machine Learning Methods based on String Kernel", International on Convergence and Hybrid Information Technology.

[28] Minqing Hu and Bing Liu, (2004), "Mining and Summarizing Customer Reviews", KDD'04, Seattle, Washington, USA, pg: 22-24.

[29] D. Lewis, M. Ringuette, (1994), "A comparison of two learning algorithms for text categorization", in: Proceedings of Third Annual Symposium on Document Analysis and IR, pg. 81–93.

[30] A. McCallum, K. Nigam, (1998), "A comparison of event models for naive Bayes text classification", in: Proceedings of AAAI-98 Workshop on Learning for Text Categorization, pg. 41–48.

[31] Qing Cao, Mark A. Thompson, Yang Yu, (2012), "Sentiment analysis in decision sciences research: An illustration to IT governance, Decision Support Systems, http://dx.doi.org/10.1016/j.dss.2012.10.026

[32] Nigam K., Lafferty J., and McCallum A, (1990), "using maximum entropy for Text Classification". In Proc of the IJCAI-99 Workshop on Machine Learning for Information Filtering.

[33] Berger A., A Brief Maximum Entropy Tutorial.

[34] Raychaudhuri S, Chang JT, Sutphin PD, and Altman RB, (2002), "Associating genes with Gene Ontology codes using a maximum entropy analysis of biomedical literature", in: Genome Research, Volume 12, pg:203-214.

[35] Manning, Christopher D. and Schutze, Hinrich, "Foundations of Statistical Natural Language Processing".

[36] Ratnaparkhi A, "A Simple Introduction to Maximum Entropy Models for Natural Language Processing".

[37] Goldman S.A., "Efficient Methods for calculating maximum entropy distributions", *MIT/LCS/TR-391.*

Special Acknowledgement

Author's short Biography

Padmaja Savaram, currently a Senior Research Fellow at the department of Computer Science and Engineering, Osmania University (India), holds an M.Tech in Computer Science and Engineering and a Masters in Mathematics from Osmania University and Pondicherry Central University respectively. She taught for over seven years and published several articles on Mobile Computing, Network Security and Embedded Systems. Her current areas of interest are Data Mining, Text Mining, Sentiment Analysis and Opinion Mining.

Dr Sameen Fatima is a BTech in Electronics and Communication Engineering from JNTU, MPhil in Computer Methods from University of Hyderabad, MS in Computer Science from the University of Massachusetts at Amherst, USA, and PhD in Computer Science and Engineering from Osmania University. She joined the Department of Computer Science and Engineering of Osmania University during its formative years in 1984. From 1989 to 1994 she worked as a Data Base Administrator at Applied Econometrics, USA and at the Physical Plant at UMASS, USA. From 2006 to 2009 she was a faculty at BITS, Pilani – Dubai campus. Currently she is Professor and Head of the Department of Computer Science and Engineering at the College of Engineering, Osmania University.