EXTRACTING THE OPINION THROUGH CUSTOMER FEEDBACK DATA FROM WEB RESOURCES

Sumeet V.Vinchurkar*

Smita M. Nirkhi**

ABSTRACT

As the customer opens their views for the product and their services, so it is necessary that it is classifying and categorizing these views in feature manner. The job of analyzing such information is collectively called as customer product reviews, also called as opinion mining. Basically it consists of some steps, and different task achieved by applying different techniques. This paper basically explains such techniques that have been used for the implementation of task of opinion mining. On the basis of this analysis we provide an overall system design for the development of opinion mining approach.

Keywords*: Opinion Mining, Customer Reviews, Blogs, Linguistic Resource, NLP Techniques, Rough Set Theory, Sentiment Classification Methods.*

*PG Department of Computer Science and Engineering, G.H. Raisoni College of Engineering, Nagpur.

**Assistant Professor, PG Department of Computer Science and Engineering, G.H. Raisoni College of Engineering, Nagpur.

I. INTRODUCTION

The text classification research in the field started with some problem related with sentiment and subjectivity classification. Sentiment classification classifies whether a feedback document (such as product reviews) or sentence expresses a positive or negative opinion. Subjectivity of classification determines whether a sentence is subjective or objective. In day to day life user wants to know about the subject opinion. For example, from a product review, users want to know which product features consumers have praised and criticized. To explore this generic problem, let's use the following review segment on Camera as an example:

"(1)I bought a Sony Cyber Shot. (2) It was such a nice camera. (3) The Picture Quality was really cool. (4) The video recording quality was clear too. (5) However, my brother was tempestuous with me as I did not tell him before I bought it. (6) He also thought the Camera was too expensive, and wanted me to return it to the store.… "

The question is, what do we want to extract from this review? The first thing that we might notice is that there are several opinions in this review.

Sentences 2, 3, and 4 express three positive opinions, while sentences 5 and 6 express negative opinions or emotions. We can also see that all the opinions are expressed about some targets or objects. For example, the opinion in sentence 2 is on the Camera as a whole, and the opinions in sentences 3 and 4 are on the Camera's picture and video recording quality, respectively. Importantly, the opinion in sentence 6 is on the Cameras's price, but the opinion/emotion in sentence 5 is about "me," not the phone. In an application, the user might be interested in opinions on certain targets but not necessarily on user-specific information. Finally, we can also see the sources or holders of opinions. The source or holder of the opinions in sentences 2, 3, and 4 is the author of the review, but in sentences 5 and 6, it is "my brother".

In this example, we can extract several phrases such as 'nice camera', 'really cool', 'video recording quality was clear', 'did not tell him', and 'the camera was too expensive', which convey customer's opinion rather than facts. In particular, subjective words such as 'nice', 'really cool', 'clear', 'did not', and 'too expensive' are used to express customer's positive/negative sentiment regarding the product features, which are referred by 'camera', 'picture', 'video'. Although information gathered from multiple reviews are more reliable compared to information from only one review, manually sorting through large amounts of review one by one requires a lot of time and cost for both businesses and customers.

Therefore it is more efficient to automatically process the various reviews and provide the necessary information in a summarized form.

Because of the importance of automatically extracting actionable knowledge from customer feedback data on the Web, "opinion mining (OM)" has become a significant subject of research in the field of data mining. The ultimate goal of OM is to extract customer opinions (feedback) on products and present the information in the most efficient way that serves the objectives chosen. This means that the necessary steps and techniques used for OM can be different depending on how the summarized information is presented. For example, if we were to get the number of negative and positive reviews about a given product, classifying each review as positive or negative would be the most important task. On the other hand, if we want to show customer feedback on each of the different features of a product, it is necessary to extract product features and analyze the overall sentiment of each feature.

The methods that are needed for feature extraction, sentiment classification, and opinion summarization have already been researched in other areas such as document classification and text summarization. These can be modified and applied to OM. However, the focus of opinion mining is on the sentiment that the customer is expressing and this is where the methods are applied differently. As can be seen from the prior example, making the linguistic distinction between objective words that express facts and subjective words that express opinions is important. In this paper, we examine the two tasks that are specific to opinion mining: development of linguistic resources and sentiment classification. In addition, we present opinion summarization by looking into the existing opinion mining systems which extract opinion expression from large reviews and show how each system applies the methods in order to effectively summarize and present opinions.

The remainder of the paper is organized as follows. Section 2 introduces the proposed architecture for opinion mining and section 3 presents the methods to be used for OM in this architecture i.e. Rough Set Theory and section 4, we discuss about sentiment classification Methods such as Fuzzy C means and Naïve Bayes with some other techniques. Section 5 concludes the paper.

II.TASK FOR OPINION MINING

As mentioned in Section 1, opinion mining can be roughly divided into three major tasks of development of linguistic resources, sentiment classification, and opinion summarization. We graphically present these tasks and the areas of research to which they are related in figure 1.

Hatzivassiloglou and Mckeown [1] are among the first to deal with opinion classification. They focus on adjectives and they study phrases where adjectives are connected with conjunction words such as «and» or «but ». They construct a log-linear regression model so as to clarify whether two adjectives have the same orientation. The accuracy of this task is declared to be 82%. The techniques used for text classification and text summarization can also be applied to OM, along with linguistic resources. Although sentiment classification and opinion summarization share several steps or techniques, sentiment classification focuses on classifying each review and opinion summarization is about how to effectively extract opinion expressions and summarize them from a large number of reviews of a given product.

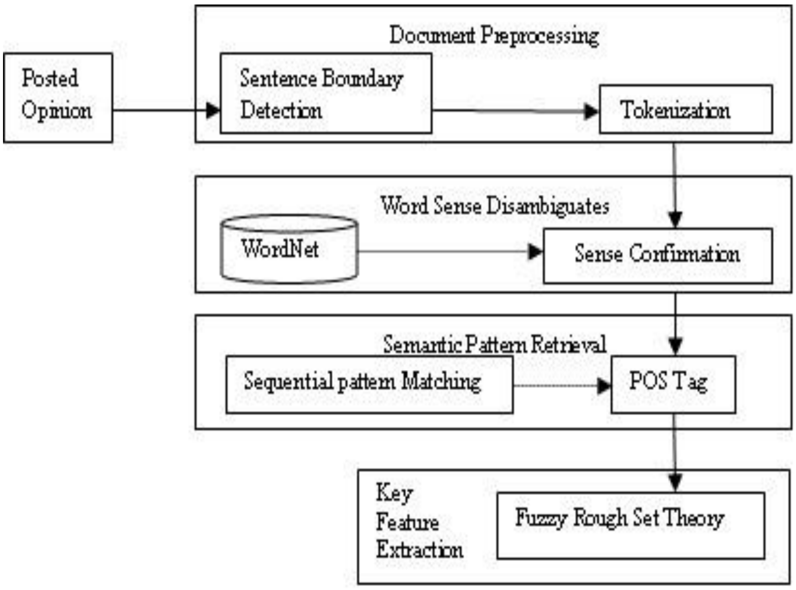

Figure1 Flow for generating opinion from feedback

Sentiment related properties are well defined in appraisal theory [2], a framework of linguistic resources for describing how writers and speakers express inter-subjective and ideological positions. However, most researches for developing linguistic resources have focused on determining three properties: subjectivity, orientation, and strength of term attitude. For example, 'good', 'excellent', and 'best' are positive terms while 'bad', 'wrong', and 'worst' are negative terms. 'Vertical', 'yellow', and 'liquid' are objective terms. 'Best' and 'worst' are more intense than 'good' and 'bad'.

2.1. Document Pre-processing

Currently, input post is of plain text format. By applying some NLP [7,15,22]steps for making the tokens is as follows.

Statement extraction*:* firstly, from each post, sentences are individuated, that are parts of text ending with a full stop, comma, question mark, exclamation mark or semicolon.

Subsequently, conjunctions are analyzed for dividing sentences into statements, which are parts of text expressing only one meaning.

Anaphora resolution*:* often, in informal text, subject and predicate could be understood; hence some statements could be incomprehensible. The goal of the anaphora resolution step is to restore a statement by adding understood parts.

Tokenization: in this step, each statement is divided into tokens, which are parts of text bounded by a separator (space, tab or end of line).

Stemming and Lemmatization*:* in order to reduce the number of different terms, each token is transformed reducing its inflectional forms to a common base form. The main difference among stemming and lemmatization's that the former extract "brutally" the root of a word (e.g. bio is the stem of biology, biocatalyst and biochemical); while the latter uses a vocabulary (often a lexical ontology) for returning the dictionary form of a word, that is the lemma (e.g. be is the lemma of is, are, was, and so on).

Tagging and Stopwords elimination*:* some word categories are too common to be useful to distinguish among statements. Hence, in this step articles, prepositions and conjunctions are first recognized and then removed. At the same time, we also remove proper nouns, which usually don't have an affective content.

2.2. Word Sense Disambiguation

In automatic text summarization, word sense disambiguation is important and many different approaches have been taken [19, 20, 22, 30]. In this process phase, the Lesk [31] approach is adopted and modified for word sense disambiguation. The Lesk approach assumes that words used in a sentence are collaborative in terms of topic and their dictionary definitions, thus must use some common words in their sense definitions. Based upon this assumption, the nouns and the verbs are first extracted from each sentence together with their senses given in WordNet [35]as the input to the following process for sense disambiguation:

1) For a word to be disambiguated, the process first scores the semantic relatedness between any two senses, one for this word and the other for any other word in the same sentence.
2) Note each sense in WordNet is semantically related with a set of similar senses. The score computed in the previous step reflects the direct relatedness between any two senses. It is like a local link. To fully reflect the relatedness between two senses, their indirect relatedness should be taken into account. The indirect semantic relatedness of two senses is the sum of the pair wise relatedness scores of the two set of similar senses.

3) The final score of each sense for the word is the sum of the score given by Step 1.) And the half value of the score given by Step 2.).
4) Among all the senses of the word, the sense with the highest score is selected as the candidate sense of the word. After all words in a sentence are disambiguated, this phase builds and reports the sense representation for the sentence in terms of WordNet senses to indicate what concept the sentence may cover.

There are three main streams in developing resources for OM: the NLP techniques, the Rough Set Theory method, and Sentiment classification method.

2.3 Natural Language Parsing Techniques

The main idea for entity discovery [5,7,19] is to discover linguistic patterns and then use the patterns to extract entity names. However, traditional methods need a large number of manually labelled training examples, and labelling is very time consuming. For a different domain, the labelling process may need to be repeated. This section proposes an automated pattern discovery method for the task, which is thus unsupervised. The basic idea of the algorithm is that the user starts with a few seed entities. The system bootstraps from them to find more entities in a set of documents (or posts). Sequential pattern mining is employed at each iteration to find more entities based on already found entities. The iterative process ends when no new entity names are found.

Pruning methods are also proposed to remove those unlikely entities. Given a set of seed entities E = {e1, e2 …en}, the algorithm consists of the following iterative steps:

Step 1 – data preparation for sequential pattern Mining

Step 2 – Sequential pattern mining

Step 3 – Pattern matching to extract candidate Entities

Step 4 – Candidate pruning

Step 5 – Pruning using brand and model relation And Syntactic patterns.

III. ROUGH SET THEORY

In Rough set theory [6,24,26,27,28], first proposed by Pawlak in 1982, employed mathematical modelling to deal with class data classification problems, and then turned out to be a very useful tool for decision support systems, especially when hybrid data, vague concepts and uncertain data were involved in the decision process. To use the rough set process, one begins with a relational database, a table of objects with attributes, and attributes values for each object.

One attribute is chosen as the decision attribute, then the rest of the attributes are the condition attributes (Pawlak, 1982). Rough sets address the continuing problem of vagueness, uncertainty and incompletion by applying the concept of equivalence classes to partition training instances according to specified criteria. Two partitions are formed in the mining process. The members of the partition can be formally described by unary set theoretic operators or by successor functions for lower approximation and upper approximation spaces from which both possible rules and certain rules can be easily derived. Vague and imprecise data sets have no clear-cut boundaries. Thus, the rough set theory approach is based on refusing certain set boundaries, implying that every set will be roughly defined using a lower and an upper approximation.

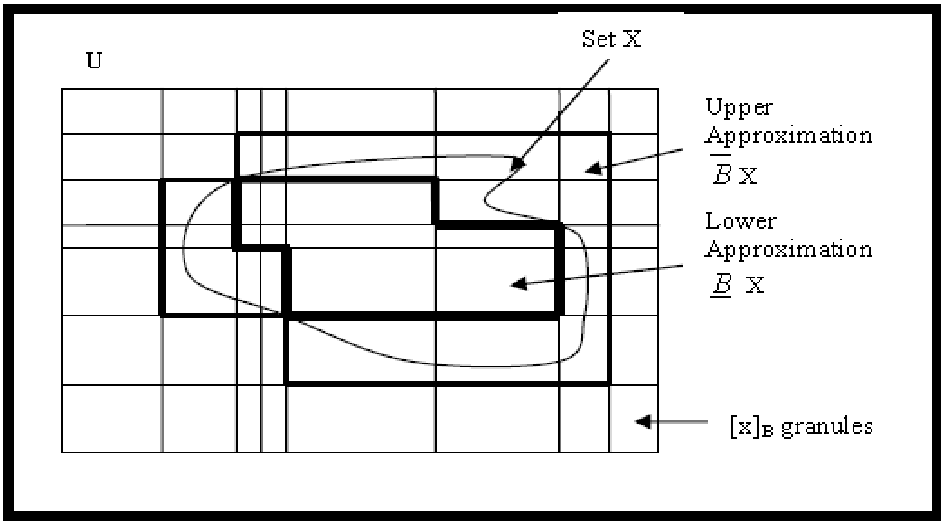

Figure 2 Rough Representation of a set with upper and lower approximations.

Let B ⊆ A and X ⊆ U be an information system. The set X is approximated using information contained in B by constructing lower and upper approximation sets:

$\underline{B}$X=$\{X|[X]_B \subseteq X\}$ (Lower Approximation)

$\overline{B}X = \{X|[X]_B \cap X \neq \emptyset\}$ (Upper Approximation)

The elements in $\underline{B}X$ can be classified as members of X by the knowledge in B. However, the elements in $\overline{B}X$ can be classified as possible members of X by the knowledge in B. The set $BN_B = \overline{B}X - \underline{B}X$ is called the B-boundary region of X and it consists of those objects that cannot be classified with certainty as members of X with the knowledge in B. The set X is called ‘‘rough” (or ‘‘roughly definable”) with respect to the knowledge in B if the boundary region is nonempty. Rough sets theoretic classifiers usually apply the concept of rough sets to reduce the number of attributes in a decision table (Pawlak, 1991) and to extract valid data from inconsistent decision

IV. SENTIMENT CLASSIFICATION METHOD

4.1 Fuzzy C Means

Fuzzy C Means Integration of fuzzy logic with data mining techniques has become one of the key constituents of soft computing in handling the challenges posed by massive collections of natural data.

The central idea in fuzzy clustering [8,26,33] is the non-unique partitioning of the data into a collection of clusters. The data points are assigned membership values for each of the clusters and the fuzzy clustering algorithms allow the clusters to grow into their natural shape. In some cases the membership value may be zero indicating that the data point is not a member of the cluster under consideration. Many crisp clustering techniques have difficulties in handling extreme outliers but fuzzy clustering algorithms tend to give them very small membership degree in surrounding clusters . The non-zero membership values, with a maximum of one, show the degree to which the data point represents a cluster. The points at the centre of the cluster have maximum membership values and the membership gradually decreases when one moves away from the cluster centre. Thus fuzzy clustering provides a flexible and robust method for handling natural data with vagueness and uncertainty. In fuzzy clustering, each data point will have an associated degree of membership for each cluster. Fuzzy C-means clustering algorithm includes two processes, the calculation of cluster centre and the assignment of points to these centres using a form of Euclidean distance. The process is continued till the cluster centre stabilizes. The algorithm incorporates the fuzzy set's concepts of partial membership and forms overlapping clusters to support it. Each data item is assigned a membership value in the range of 0 to 1 for the clusters. Degree of fuzziness in the clusters is indicated by the parameter named as fuzzification (m). When the value of m is equal to 1 the algorithm works like a crisp partitioning algorithm and for larger values overlapping of cluster tends to be more. Membership of each data item is calculated using (1).

$$\mu_j (x_i)=\frac{\left(\frac{1}{d_{ji}}\right)^{\frac{1}{m-1}}}{\sum_{k=1}^{p}\left(\frac{1}{d_{ki}}\right)^{\frac{1}{m-1}}} \qquad (1)$$

Where $\mu_j(x_i)$- indicates the membership of x_i in the j^{th} cluster

d_{ji} – distance of x_i in cluster c_j

m – Fuzzification parameter

p – Number of specified clusters

d_{ki} – distance of x_i in cluster c_k

The sum of memberships of a data point in all clusters must be equal to 1. The new cluster centers are calculated using (2)

$$C_j = \frac{\sum_i \mu_j (x_i)^{m(xi)}}{\sum_i \mu_j (x_i)^m} \qquad (2)$$

The algorithm begins by choosing the number of clusters and fuzzification parameter. Centre for all the clusters are chosen randomly. The algorithm continues to update the centre of the clusters till the value stabilizes.

4.2 Naive Bayesian classifiers

Let C denote a class attribute with a finite domain dom(C) of m classes, and V_1 . . . $V_{n.}$ a number at-tributes with finite domains dom(V_1). . .dom(Vn). An example e is described by its attribute values $v\ {}^e_1 \in dom(v_{1)}$,...... $v\ {}^e_n \in dom(v_{n)}$.To simplify the presentation, we assume in the remainder of the pa-per that all of these values actually occur in the training examples, such that none of the discussed probabilities are zero. In practice, this demand can easily be fulfilled by only considering the values that al-ready occurred in the examples. Naive Bayes classifiers[7,33,34] implement a probabilistic

idea of classification: they calculate the class of a new example by estimating for each class from

dom(C) the probability that the example is in this class, and predict the most probable class as the class ofe. Formally, for all C ∈ dom(C). they estimate the probability $P\big((C = C|V_1 = v_1^e, V_2 = v_2^e, .., V_n = v_n^e)\big)$ that an example with attribute values like the given

new example has the class c. To improve read-ability, we will use in the following P(. . v_1^e. . .).as an abbreviation for P(. . . _$V_i = v_1^e$. . .), as well asP(. . . c . . .) as an abbreviation for P(. . . C =c . . .).In practice, a complete table of these conditional probabilities would usually be too large to memorize, because its size is exponential in the number of attributes. Moreover, the number of training examples is often too small to give a good estimate for all these values. So, the idea of Naive Bayes classification is first to apply the Bayes rule

$$\boldsymbol{P}(\boldsymbol{Y}|\boldsymbol{X}) = \frac{\boldsymbol{P}(\boldsymbol{X}|\boldsymbol{Y})\boldsymbol{P}(\boldsymbol{Y})}{\boldsymbol{P}(\boldsymbol{X})}$$

The independence assumption is naive in that it is in general not met. Nevertheless, Naive Bayesian classifiers give quite good results in many cases, and are often a good way to perform classification when there is too little data on those dependencies to employ more sophisticated means. As [1] discusses, they can be seen as a special case of Bayesian Networks [6], which take attribute dependencies into consideration.

4.3 Some Other algorithms

Sentiment classification is the process of identifying the sentiment - or polarity of a piece of text or a document. In this section, we discuss certain clustering algorithms [3,6,11] for polarity classification.

C4.5 and beyond [3,24] deals with the Systems that construct classifiers are one of the commonly used tools in data mining. Such systems take as input a collection of cases, each belonging to one of a small number of classes and described by its values for a fixed set of attributes, and output a classifier that can accurately predict the class to which a new case belongs.

The k-means algorithm [3,6] is a simple iterative method to partition a given dataset into a user specified number of clusters, k. The algorithm operates on a set of d-dimensional vectors, $D = \{\frac{x}{i} = 1 \ldots\ldots N\}$, where $\mathbf{x}\mathrm{i} \in \Re\mathrm{d}$ denotes the i th data point. The algorithm is initialized by picking k points in $\Re d$ as the initial k cluster representatives or "centroids". Techniques for selecting these initial seeds include sampling at random from the dataset, setting them as the solution of clustering a small subset of the data or perturbing the global mean of the data k times. Then the algorithm iterates between two steps till convergence:

Step 1: Data Assignment.

Step 2: Relocation of "means".

In today's machine learning applications, ***Support Vector Machines*** **(SVM)** [3,6] are considered a must try—it offers one of the most robust and accurate methods among all well-known algorithms. It has a sound theoretical foundation, requires only a dozen examples for training, and is insensitive to the number of dimensions. In addition, efficient methods for training SVM are also being developed at a fast pace. In a two-class learning task, the aim of SVM is to find the best classification function to distinguish between members of the two classes in the training data. The metric for the concept of the "best" classification function can be realized geometrically. For a linearly separable dataset, a linear classification function corresponds to a separating hyper plane $f(x)$ that passes through the middle of the two classes, separating the two. Once this function is determined, new data instance x^n can be classified by simply testing the sign of the function $f(x^n)$; x^n belongs to the positive class if $f(x^n) > 0$. Because there are many such linear hyper planes, what SVM additionally guarantee is that the best such function is found by maximizing the margin between the two classes. Intuitively, the margin is defined as the amount of space, or separation between the two classes as defined by the hyper plane. Geometrically, the margin corresponds to the shortest distance between the closest data points to a point on the hyper plane. Having this geometric definition allows us to explore how to maximize the margin, so that even though there are an infinite number of hyper planes, only a few qualify as the solution to SVM.

Table1. Classification algorithms with its characteristics

Algorithm	**Input to the system**	**Characteristics**	**Way of Generating**	**Ongoing Research Work**
C4.5 and beyond	Collection of cases	Ready to show two or more outcomes	Classifier shows the classification of new case	Variant of boosting and Misclassification cost
k-means	Sample from dataset	Iterative model	Centroid of clusters	Whenever the data is not describing reasonably
support vector machines	Trained dataset	Best Accuracy	Hyper planes	Rule based
Apriori	Itemset	Not under machine learning	Association rule for frequent itemset	Eliminating candidate generation
Naive Bayes	Set of objects	Supervised classification	Probability of objects	Improvement in predictive accuracy

Apriori [3,6] is a seminal algorithm for finding frequent itemsets using candidate generation .It is characterized as a level-wise complete search algorithm using anti-monotonicity of itemsets, “if an itemset is not frequent, any of its superset is never frequent”. By convention, Apriori assumes that items within a transaction or itemset are sorted in lexicographic order. Let the set of frequent itemsets of size k be F_k and their candidates be C_k. Apriori first scans the database and searches for frequent itemsets of size 1 by accumulating the count for each item and collecting those that satisfy the minimum support requirement. It then iterates on the following three steps and extracts all the frequent itemsets. Many of the patterns finding algorithms such as decision tree, classification rules and clustering techniques that are frequently used in data mining have been developed in machine learning research community. Frequent pattern and association rule mining is one of the few exceptions to this tradition. The introduction of this technique boosted data mining research and its impact is tremendous. The algorithm is quite simple and easy to implement. Experimenting with Apriori-like algorithm is the first thing that data miners try to do. Table 1 shows the characteristics of algorithms.

V. DISCUSSION

We provided an overall picture of the tasks and techniques involved in developing an automated system for mining opinions that are found in customer feedback data on the Web. More specifically, we were able to develop this overall picture by conducting a survey and analysis of the techniques involved in each of the key steps of opinion mining. In this paper, we focused on surveying and analyzing the methods for development of linguistic resources, sentiment classification. Even now, on the Web, vast amounts of user generated contents are being created on merchant sites and internet forums. These contents have been recognized as measurable resources and various opinion mining methods have been developed to analyze the contents. Furthermore, opinion mining has become important for all types of organizations, including for-profit corporations, government agencies, educational institutions, non-profit organizations, the military, etc. to gauge the opinions, likes and dislikes, and the intensity of the likes and dislikes, of the products, services, and policies they offer and plan to offer. As such, we believe that an understanding of the overall picture of the tasks and Techniques involved in opinion mining is of significant importance.

REFERENCES

1. Hatzivassiloglou, V. and K. Mckeown (1997). Predicting the semantic orientation of adjectives. In Proceedings of the 8th conference on European chapter of the Association for Computational Linguistics, pp. 174–181.
2. Martin, J.R., White Peter R. R, the Language of Evaluation: Appraisal in English. 2005: Palgrave Macmillan.
3. Han, J., and Kamber M. Data Mining: Concepts and Techniques, 2006
4. Nitin Jindal and Bing Liu," Mining Comparative Sentences and Relations" in American Association for Artificial Intelligence, AAAI-2006

5. Xiaowen Ding, Bing Liu, Lei Zhang in "Entity Discovery and Assignment for Opinion Mining Applications". KDD'09, ACM, June 28– July 1, 2009, Paris, France.
6. XindongWu · Vipin Kumar · J. Ross Quinlan · Joydeep Ghosh · Qiang Yang ·Hiroshi Motoda · Geoffrey J. McLachlan · Angus Ng · Bing Liu · Philip S. Yu · Zhi-Hua Zhou · Michael Steinbach · David J. Hand · Dan Steinberg, "Top 10 algorithms in data mining" in Springer Knowl Inf Syst (2008) 14:1–37
7. Durga Toshniwal and Rishiraj Saha Roy:"Clustering Unstructured Text Documents Using Naïve Bayesian Concept and Shape Pattern Matching" in International Journal of Advancements in Computing Technology Volume 1, Number 1, September 2009, pp.52-63
8. K.Suresh, R.Madana Mohana, A.Rama Mohan Reddy," Improved FCM algorithm for Clustering on Web Usage Mining" in IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 1, January 2011,pp.42-46
9. Kushal Dave, Steve Lawrence, David M. Pennock: Mining the peanut gallery: opinion extraction and semantic classification of product reviews. WWW 2003: 519-528
10. Minqing Hu, Bing Liu: Mining Opinion Features in Customer Reviews. To appear in AAAI'04, 2004.
11. N. Sivaram, K. Ramar: Applicability of Clustering and Classification Algorithms for Recruitment Data Mining in International Journal of Computer Applications (0975 – 8887) Volume 4 – No.5, July 2010:23-28.
12. Hiroshi Kanayama, Tetsuya Nasukawa, & Hideo Watanabe:Deeper sentiment analysis using machine transalation technology. Coling 2004: 20th International Conference on Computational Linguistics, 23-27 August 2004, University of Geneva, Switzerland, Proceedings.
13. Ana-Maria Popescu, Oren Etzioni: Extracting Product Features and Opinions from Reviews. HLT/EMNLP 2005
14. Lingyan Ji, Hanxiao Shi, Mengli Li, Mengxia Cai, Peiqi Feng. Opinion Mining of Product Revies Based on Semantic Role Labeling in The 5th International Conference on Computer Science Education Hefei, China. August 24–27, 2010,pp.1450-1453.
15. Domenico Potena, Claudia Diamantini. Mining Opinions on the Basis of Their Affectivity in 2010 IEEE, pp.245-254.
16. Wei Jin and Hung Hay Ho. A Novel Lexicalized HMM- based Learning framework for Web Opinion Mining in Proceedings of the 26th International Conference on Machine Learning, Montreal, Canada, 2009.

17. Themis P. Exarchos · Markos G. Tsipouras ·Costas Papaloukas · Dimitrios I. Fotiadis .An optimized sequential pattern matching methodology for sequence classification in Knowl Inf Syst, 12 April 2008© Springer-Verlag London Limited 2008
18. Nitin Jindal and Bing Liu Identifying Comparative Sentences in Text Documents SIGIR'06, August 6-11, 2006, Seattle, Washington, USA.
19. Xiaowen Ding, Bing Liu and Philip S. Yu A Holistic Lexicon-Based Approach to Opinion Mining WSDM'08, February 11-12, 2008, Palo Alto, California, USA
20. Minqing Hu and Bing Liu Mining and Summarizing Customer Reviews KDD'04, August 22–25, 2004, Seattle, Washington, USA.
21. Murthy Ganapathibhotla and Bing Liu Mining Opinions in Comparative Sentences © 2008IEEE
22. Qiuye Zhao and Mitch Marcus. A Simple Unsupervised Learner for POS Disambiguation Rules Given Only a Minimal Lexicon in Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, pages 688–697,Singapore, 6-7 August 2009.
23. Asif Ekbal and Sivaji Bandyopadhyay. Web-based Bengali News Corpus for Lexicon Development and POS Tagging in SPSAL-Proceedings June12, 2008.
24. Ching-Hsue Cheng,You-Shyang Chen," Classifying the Segmentation of customer value via RFM model and RS Theory" in journal Expert system with application 36 (2009),pp.4176-4184
25. Yiyuan Cheng, Ruiling Zhang, Xiufeng Wang, Qiushuang Chen. Text Feature Extraction Based on Rough Set in Fifth International Conference on Fuzzy Systems and Knowledge Discovery, 2008 IEEE, pp.310-314.
26. Hsun-Hui Huang, Yau-Hwang Kuo and Horng-Chang Yang. Fuzzy-Rough Set Aided Sentence Extraction Summarization in Proceedings of the First International Conference on Innovative Computing, Information and Control (ICICIC'06).
27. Qiang Li, Jjan-Hua Li, Gong- Shen liu, Sheng-Hong Li. A Rough Set based Hybrid Feature Selection Method For Topic Specific Text Filtering in Proceedings of the Third International Conference on Machine Learning and Cybernetics, Shanghai, 26-29 August 2004, pp.1464-1468.
28. Richard Jensen and Qiang Shen . Semantics-Preserving Dimensionality eduction: Rough and Fuzzy-Rough-Based Approaches in IEEE Transactions on Knowledge and Data Engineering, Vol. 16, No. 12, December 2004,pp.1457-1471.

29. Zdzisław Pawlak. Rough set theory and its applications in journal of Telecommunication and Information Technology 3/2002, pp.7-10.
30. Siddharth Patwardhan, Satanjeev Banerjee, and Ted Pedersen, "Using Measures of Semantic Relatedness for Word Sense Disambiguation" In Proceedings of the Fourth International Conference on Intelligent Text Processing and Computational Linguistics, Mexico City, 2003.
31. Satanjeev Banerjee and Ted Pedersen, "An Adapted Lesk Algorithm for Word Sense Disambiguation Using WordNet" In Proceedings of the Third International Conference on Intelligent Text Processing and Computational Linguistics,Mexico City, pp. 136–145, 2002.
32. Hans-Peter Storr. A compact fuzzy extension of the Naive Bayesian classification algorithm. AI Institute, Dept. Computer Science Technische Universit¨ at Dresden
33. Yongchuan Tang, Wuming Pan, Haiming Li and Yang Xu." Fuzzy Naive Bayes Classifier Based on Fuzzy Clustering" In 2002 IEEE .
34. Miller, G., Beckwith, R, Fellbaum, C., Gross, D., and Miler, K. 1990. Introduction to WordNet: An on-line lexical database. International Journal of Lexicography (special issue), 3(4):235-312.
35. Dongjoo Lee, Ok-Ran Jeong, Sang-goo Lee. Opinion Mining of Customer Feedback Data
on the Web in Ministry of Information & Communications, Korea, under the Information Technology Research Center (ITRC) Support Program.