Enhanced Discoverability of Content through Linked Data for Online Reviews using Classification and Ranking Techniques

B. Dhanalakshmi
Sathyabama University, Chennai

A.Chandrasekar
St Joseph's College of Engineering, Chennai

ABSTRACT
Massive unstructured data are available and being posted in numerous blogs, forums, and online sites. This enormous amount of information on worldwide network platforms make them feasible and can be used as input source, in applications based on opinion mining and sentiment analysis. The aim of this paper is to analyze online reviews in unstructured form and discover content through linked data and deriving an opinion. Our proposed methodology comprises of phases such as Data pre-processing, content discovery and Opinion mining. Initially the unstructured data is extracted from the web document. This phase is used for formatting the data before sentiment analysis and mining. The second phase will be classified into two i.e., Feature extraction, content discovery and opinion extraction. The features like term frequency, Part of Speech are extracted from the words in the documents. After feature extraction, we extract useful information related to the item's features and use it to rate them as positive, neutral, or negative. This final phase will be done by supervised learning algorithm decision tree classifier with the help of features extracted. In the final step ranking and classification will be done.

Keywords
Opinion Mining, Sentiment Analysis, Content Discovery, Opinion Extraction.

1. INTRODUCTION
Opinion mining deals with the sentimental word which clearly demonstrates ideas about the targeted object at that particular time. It is a type of natural language processing for gathering the information from public about a particular product. Whenever we decide to buy something new, we always look forward for someone's advice for same. This field has wide range of application due to its increasing popularity among the public. For example, it can help you judge that which product is good and which is bad while both are having the same price, market analysts may use the reviews posted in various blogs for taking major decisions regarding new product launch. Opinion mining actually identifies the author's viewpoint about a subject, rather than simply identifying the subject itself [1]. Opinion mining aims at recognizing and categorizing or extracting opinions found in unstructured text resources and are one of the most dynamically. Main objective is to classify an opinion according to a polar spectrum. Opinions are the views that are expressed by an individual on some topic/issue according to his/her own perspective.

As these views have been used by several application domains such as business and organization, individual etc., we can say it has become very important to find out efficient ways for extracting opinions. So Opinion mining is the process of studying people's opinions or emotions towards entities, events and their features. In the past few years, it has attracted a great deal of attentions from both academia and industry due to many challenging research problems and a wide range of applications. Another name for opinion mining may be sentiment analysis and subjective analysis. The objective of opinion analysis is to identify emotions from text and determine their polarities. On the basis of determined polarities we can conclude whether the text document represents positive ,negative or neutral opinion .The discovered opinions are useful to many practical applications such as opinions in product quality reviews are helpful to potential customers. Users also comment on products in their personal web sites and blogs, which are then aggregated by sites such as Blogstreet.com, AllConsuming.net, and onfocus.com [2]. Meanwhile, opinion mining technique relates indirectly to promote many natural language processing techniques. The sudden growth in the area of opinion mining, which deals with the computational techniques for opinion extraction and understanding created an utmost need to understand and view the internet in a different prospect [3].

With the successful adoption of Web 2.0, the Web has dramatically changed the that people share opinions and views. The wide coverage of topics and wealth of opinions make the user-generated content an extremely valuable source for mining opinions concerning products, services [4].In business and trade, these electronic word-of-mouth (eWOM) communication impacts online retailers as this user-generated content could greatly affect the decision of consumption. According to a survey, 81% of online users have read online product reviews before they go shopping and about 73% users think that online reviews play an important role in making heir purchase decisions [5]. User-generated content can also provide companies with abundant opinion information for making marketing strategies, boosting reputation and managing public relations. Unfortunately, among a great many of online reviews, it is extremely difficult for users to read and make informed purchase decisions.

A seller's job can be quite complicated or should I say it can be quite easy. The two contradictory terms define the selling experience, based on the fact as how seller interprets the consumer interests. Unless you are psychic or you know to get into others mind you can't collaborate the actual demands of the consumer and what the product is going to supply. Worse still, many reviews are redundant and irrelevant, which increase the difficulty for a potential customer to use. Besides, the large amount of reviews makes it hard for companies to keep track of public opinions on their products. Thus, it is crucial to mine and summarize opinions from online reviews automatically [6][7]. In online reviews, people usually express their opinions on product features which include components, attributes, and functions. Moreover, a positive opinion on a product feature does not mean that the user likes everything about the product and vice versa. Consequently, opinion mining against consumers' reviews is often performed at the fine grained level to provide deep analysis for the target product [6][8][9]. Opinions play important role in the process of knowledge discovery or information retrieval and can be considered as a sub discipline of Data Mining. A major interest has been shown towards the automatic extraction of human opinions from web documents in the following sections.

2. RELATED WORK

Tanvir Ahmad and Mohammad NajmudDoja [10] proposed a rank based system for features from user generated contents for different models of camera. They firstly identified the features and their modifiers and then found there polarity values. Secondly they calculated the weight of each features and ranked them on the basis of their score values. They have also separated the positive and negative features so that the user and the manufacturer would know the features which are generally liked and disliked by the user. Manufacturer accordingly developed business plans so that necessary improvement can be done in those areas. It observed that the recall values of the system was low since a sizeable amount of reviewers did not use correct English and the parser fails to identify the sentence and did not give correct POS.

Qingliang Miao et al. [11] investigated three basic sub-tasks of fine grained opinion mining and merge product feature and opinion extraction in a unified process by using CRFs models. As part of their work, they have designed a computational approach to acquire and exploit domain specific sentiment lexicon. Experimental results indicated that, their domain specific lexicon was efficient, especially for domain sensitive opinions. They also confirmed that feature cluster indeed reduce the redundancy. In addition, theyhave find the effectiveness of feature cluster varies from domain to domain. The experimental results indicatedtheir sentiment lexicon was efficient to identify the polarity of opinions, especially on domain sensitive opinion words. As a future research, theyplanned to incorporate more contextual clues that cross sentence boundaries. They also planned to develop a bootstrap approach which can make more use of unlabeled reviews, and adopt pronoun resolutions in our models as well. First, they treated the task of product feature and opinion extraction as sequence labeling process, and adopt a probabilistic model to merge product feature and opinion extraction in a unified process. Second, they developed a computational approach to construct domain specific sentiment lexicon based on general sentiment lexicon and semi-structured reviews, which used later to assist opinion polarity identification and product feature clustering.

Ion Smeureanu and CristianBucur [12] presented The expression of opinions of users in specialized sites for evaluation of products and services, and also on social networking platforms, has become one of the main ways of communication, due to spectacular development of web environment in recent years. The large amount of information on these platforms made them viable for use as data sources, in applications based on opinion mining and sentiment analysis. Proposed paper presented a method of sentiment analysis, on the review made by users to movies. Classification of reviews in both positive and negative classes done based on a naive Bayes algorithm. As training data they used a collection (pre-classified in positive and negative) of sentences taken from the movie reviews. To improve classification they removed insignificant words and introduced in classification groups of words (n-grams). For n = 2 groups they achieved a substantial improvement in classification. As an extension of the research presented in proposed paper they want to improve the algorithm, enriching the training set of examples, on the way, with examples classified as strong positive or negative, by an established score of classification. They tried to determine, in a review, those sentences which did not express opinions, or determine opinions about the film or the film actors and identify opinion addressed strictly on these items. They tried to highlight the main aspects on which opinions expressed and to extract opinions based on aspects identification.

Hsinchun Chen [13] presented an approach,which was intended to illustrate how IFS can be combined with larger feature sets for enhanced opinion-classification performanceTherewere many ways in which IFS for opinion classification can be extended in future research. Numerous additional feature categories could be used, resulting in even more robust feature sets. The syntactic and semantic information modules could be expounded on, for instance, by incorporating additional lexical resources and real-world knowledge bases. Traditionally, sentiment-analysis research has relied on two types of feature occurrence measures (frequency and presence), while researchers have yet to methodically explore additional distributional and positional measurements. Distributional measured such as compactness and first appearance have been successfully applied to topic-based text categorization. These measures used to supplement existing occurrence measures. Hence, they used IFS mechanisms to reduce opinion- classification feature spaces in a 2D manner: across feature categories (such as specific text features) and various occurrence measures associated with those features. Future feature-selection efforts explored the unique challenges associated with performing opinion classification at the document-level versus sentence, phrase, and word-level classification. Furthermore, there were other sentiment-analysis tasks that could benefit from improved feature selection, such as opinion holder identification and sentiment target detection.

G.Vinodhini and RM.Chandrasekaran [14] presented Sentiment detection has a wide variety of applications in information systems, including classifying reviews, summarizing review and other real time applications. There were likely to be many other applications that were not discussed. It found that sentiment classifiers were severely dependent on domains or topics. It was evident that neither classification model consistently outperforms the other,different types of features have distinct distributions. It was also found that different types of features and classification algorithms combined in an efficient way in order to overcome their individual drawbacks and benefit from each other's, and finally enhanced the sentiment classification performance. In future, more work needed on further improving the performance measures. Sentiment analysis 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 expected 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 dedicated to these challenges.

Qi Su1 et al.[15] proposed a novel algorithm to deal with the feature-level product opinion mining problem. An unsupervised approach based on the mutual reinforcement principle proposed. The approached clusters product features and opinion words simultaneously and iteratively by fusing both their content information and link information. Based on the clusterings of the two interrelated data objects, they constructed an association set between product feature categories and opinion word groups by identifying the strongest n sentiment links. Thus they exploited the sentiment association hidden in reviews. Moreover, knowledge from multi-source used to enhance clustering in the procedure. Their approach largely predictedopinions relating to different product features, even for the case without explicit appearance of product feature words in reviews. The experimental results based on real Chinese web reviews demonstrate that their method outperforms the state-of-art algorithms. Although their methods of candidate product feature extraction and filtering partly identify real product features, it loosed some data and remains some noises.

LI Caiqiang and Ma Junming [16] proposed a new teacher evaluation method based on opinion mining named OMTEM, considering the students' online comments in the LMS. The web 2.0 application was widely used. Combining the students' comments, OMTEM was direct, useful, and persuasive in teacher evaluation. It made up the shortcomings of firmed index in traditional teacher evaluation. However, further work need be carried out on characteristic value recognition and polarity word dictionary in order to achieve a better result.The proposed paper provided an online education teacher evaluation model based on opinion mining. The model collected opinion texts in the LMS by using web crawler, and processes them by using topic extraction and sentiment orientation classification, etc.

3. PROBLEM DEFINITION

In recent years, the spectacular development of web technologies, lead to an enormous quantity of user generated information in online systems. This large amount of information on web platforms make them viable for use as data sources, in applications based on opinion mining and sentiment analysis [12]. The development of internet and web 2.0 technologies, enabled by cost reduction of technological infrastructure, has been an exponential increase in the amount of information in online systems. These very large volumes of information are very difficult to process by individuals, leading to information overload and affecting decision-making processes in organizations. Therefore, providing new techniques for creation of knowledge is important in organizational strategy [17].

The sole purpose of Sentiment Analysis is to facilitate online consumers in decision making process of purchasing new products. Opinion Mining deals with searching of sentiments that are expressed by Individuals through on-line reviews, surveys, feedback, personal blogs etc. With the vast increase in the utilization of Internet in today's era a similar increase has been seen in the use of blog's, reviews etc. The person who actually uses these reviews or blog's is mostly a consumer or a manufacturer. As most of the customers of the world are buying & selling product on-line so it becomes company's responsibility to make their product updated. In the current scenario companies are taking product reviews from the customers and on the basis of product reviews they are able to know in which they are lacking or strong this can be accomplished with the help of sentiment analysis [3].

The explosive growth of the user-generated content on the Web has offered a rich data source for mining opinions. However, the large number of diverse review sources challenges the individual users and organizations on how to use the opinion information effectively. Therefore, automated opinion mining and summarization techniques have become increasingly important [11].

4. PROPOSED METHODOLOGY

The primary intension of this research is to develop a technique for extracting the opinions from the online user reviews. The proposed methodology comprises of 3 major phases such as 1) Data Preprocessing 2) Opinion extraction3) Opinion mining. Initially the data extracted from the web document which is unstructured. This phase is used for formatting the data before sentiment analysis and mining. The second phase will be classified into two i.e., Feature extraction and opinion extraction. The features like term frequency, Part of Speech (POS) are extracted from the words in the documents. After feature extraction, we extract useful information related to the item's features and use it to rate them as positive, neutral, or negative. This final phase will be done by supervised learning algorithm naïve bayes classifier or decision tree classifier with the help of features extracted. In the final step ranking and classification will be done. The implementation will be done in JAVA.

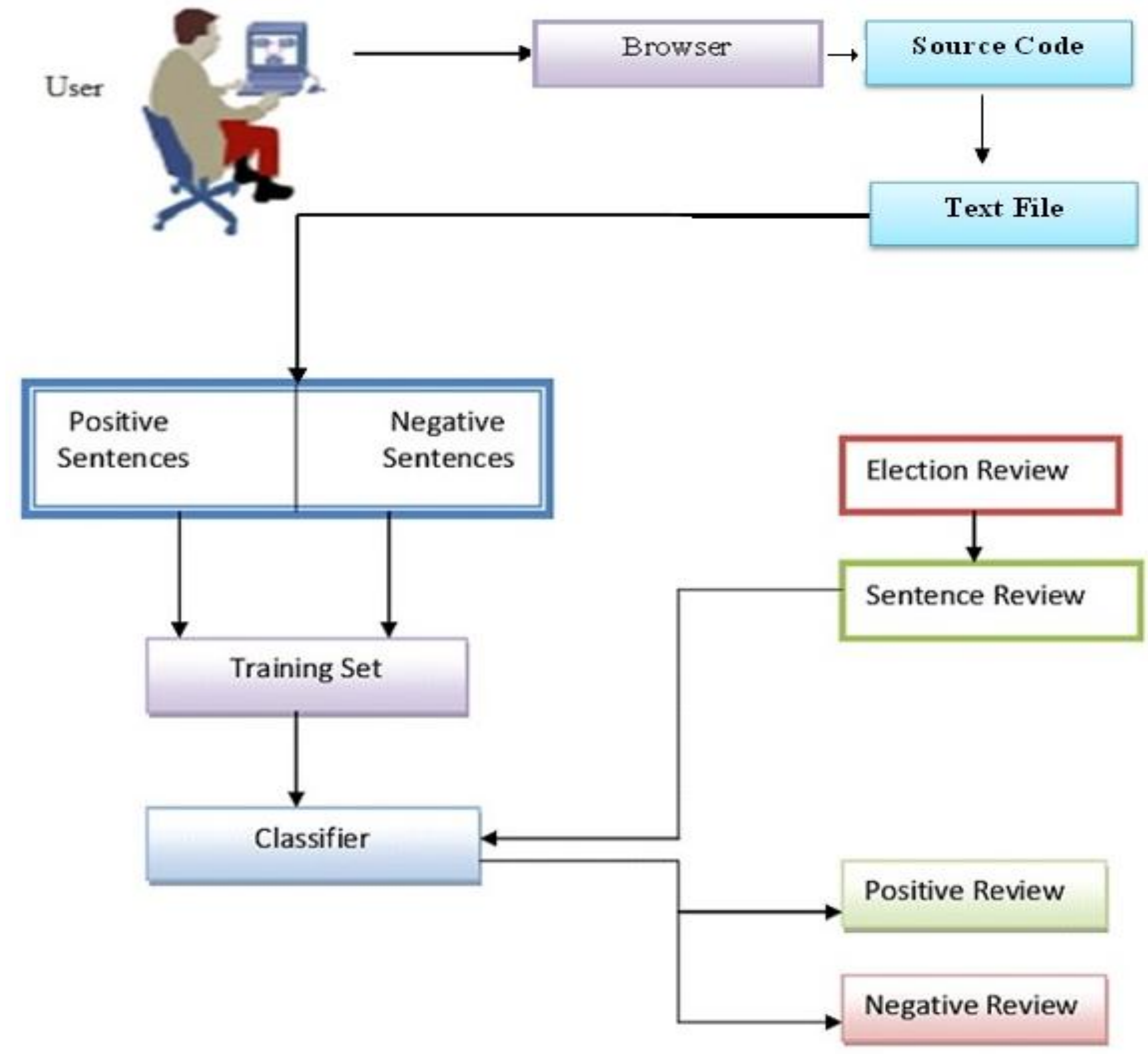

Fig. 1: Stages of classification process

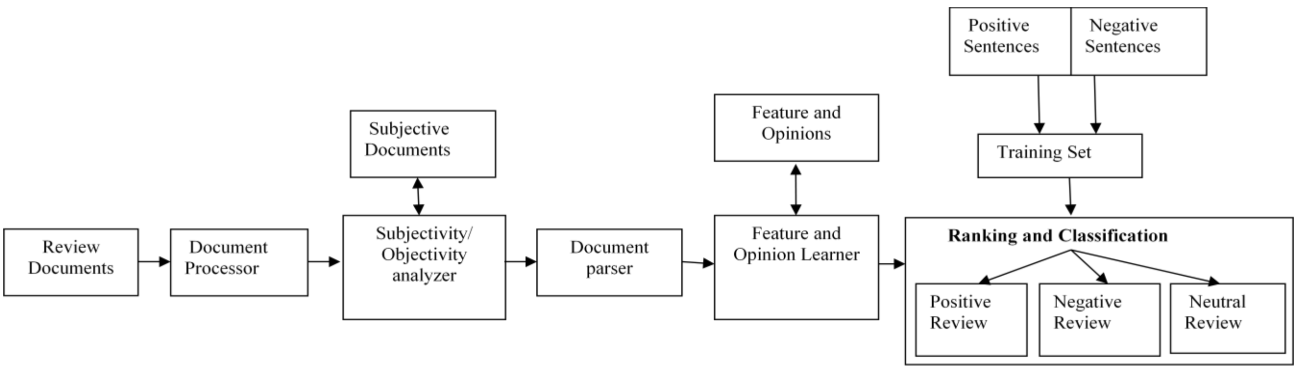

Fig. 2: Architecture of Opinion Mining (Flow of Proposed Approach)

5. DATA EXTRACTION FROM WEB DOCUMENT

Initially the data extracted from the web document is given as input and opinions are analyzed based on the user reviews as given in Fig. 1. The data extraction from the web document is the complex process. Data extraction is the act or process of retrieving data out of (usually unstructured or poorly structured) data sources for further data processing or data storage (data migration). The import into the intermediate extracting system is thus usually followed by data transformation and possibly the addition of metadata prior to export to another stage in the data workflow. Usually, the term data extraction is applied when (experimental) data is first imported into a computer from primary sources, like measuring or recording devices. Today's electronic devices will usually present an electrical connector (e.g. USB) through which 'raw data' can be streamed into a personal computer.

Then sentiment analysis is done. The opinions are analyzed. Sentiment analysis is performed in three stages after that data mining is done [18].

5.1 Sentiment Analysis

Sentiment analysis has many applications. It was done in elections to analyze voter's opinions. It was done in financial markets to identify the price trends. Social networks like face book and Twitter play a major role in sentiment analysis. Public can express their sentiments in such social forums. There are three important stages in sentiment analysis. They are

- Data collection and pre-processing
- Classification
- Aggregation and presentation of results

5.1.1 Data collection and pre-processing

Data collected from the Web document and the elimination of all matters except the opinions was done. Opinions are detected and were obtained from the web document. Pre-processing was done to eliminate all the unnecessary words or irrelevant opinions. Keywords should be extracted from the text that can provide an accurate classification. The storage of keywords should be done as the array of features A, as shown in equation (1).

$$A=(A1, A2,......, An) \quad \textbf{(1)}$$

Each element of the array is called as the feature which is a word from the original text. There may be a binary value present for every feature or a value may exist which can express the frequency of appearance in the text. Selecting common aspects is necessary.

5.1.2 Classification

Classification was done to identify the content polarity. Mainly three classes were used for classification. They are Positive, Negative and Neutral as given in Fig. 2. Classification algorithm shown below is used for sentiment analysis, depending on the method used, requires a training set with examples. It is important to train the model used for classification with specific data. Marking is done by expressing subjectivity and polarity of training sets The Classification algorithm has following steps:

Initialize P (pos) <- nr_popozitii (pos) / nr_total_propozitii

Initialize P (neg) <- nr_popozitii (neg) / nr_total_propozitii

Tokenize sentence in words

For each class of {pos, neg}:

For each word in {phrase}

P (word | class) <- nr_apartii (word | class) 1 / nr_cuv class) + nr_total_cuvinte

P (class) <-P (class) * P (word | class)

Returns max {P(pos), P(neg)}

End For

5.1.3 Aggregation and presentation of results

Aggregation is done to determine the general opinion of the analyzed text. Presentation can be done directly expressing sentiment textual or using graphics.

After the application of sentiment analysis to the input data is extracted from the web document, the next step is feature extraction.

5.2 Feature Extraction

The features extracted from the text are

- Term presence vs. frequency
- Parts of speech

5.2.1 Term presence vs. frequency

It is traditional in information retrieval to represent a piece of text as a feature vector wherein the entries correspond to individual terms. One influential finding in the sentiment-analysis area is as follows. Term frequencies have traditionally been important in standard IR, as the popularity of tf-idf weighting shows; but in contrast, obtained better performance using presence rather than frequency. That is, binary-valued feature vectors in which the entries merely indicate whether a term occurs (value 1) or not (value 0) formed a more effective basis for review polarity classification than did real-valued feature vectors in which entry values increase with the occurrence frequency of the corresponding term.

5.2.2 Parts of speech

Part-of-speech (POS) information is commonly exploited in sentiment analysi

s and opinion mining. One simple reason holds for general textual analysis, not just opinion mining: part-of-speech tagging can be considered to be a crude form of word sense disambiguation. Adjectives have been employed as features by a number of researchers. The fact that adjectives are good predictors of a sentence being subjective does not, however, imply that other parts of speech do not contribute to expressions of opinion or sentiment. After the extraction of the features such as Term presence vs. frequency and Parts of speech, opinion mining is done. Opinion mining is done with the help of Naive Bayes algorithm. The Naive Bayes classifier is a probability classifier, based on Bayes' theorem.

5.3 Feature Extraction

Fig. 2 shows the architecture of Opinion mining using Decision Tree algorithm. The complete framework of the Opinion Mining of feature words consists of the following five major modules.

- Document Processor
- Subjectivity/Objectivity Analyzer
- Document Parser
- Feature and Opinion Learner
- Feature Visualizer and Ranker

5.3.1 Document Processor

Document Processor module is used for identifying relevant portion of the text documents. Markup Language tag filter was used here which divides the individual documents in individual record size chunks and presents them as individual unstructured record documents for further processing. The cleaned document is then converted into numeric-vectors using unigram model for the purpose of subjectivity/objectivity analysis.

5.3.2 Subjectivity/Objectivity Analyzer

Subjectivity analysis is used to retain segments (sentences) of a review that are more subjective in nature and filter out those that are more objective. This increases the system performance both in terms of efficiency and accuracy. This is used to divide the reviews into subjective parts and objective parts. Cohesiveness is used to indicate segments of a review that are more subjective in nature versus those that are more objective. The training set is used to get the probability for each word to be subjective or objective, and the probability of a sentence to be subjective or objective is calculated using the unigram model. The training set is used to train the decision tree classifier.

5.3.3 Document Parser

Document parser assigns Parts-Of-Speech (POS) tags to English words based on the context in which they appear. The POS information is used to locate different types of information of interest inside the text documents. The dependency tree, also known as word-word relationship, encodes the grammatical relations between every pair of words.

5.3.4 Feature and Opinion Learner

This module is used to extract feasible information component from review documents. Information components are processed to identify product features and opinions. It takes the dependency tree generated by Document Parser as input and output the feasible information component after analyzing noun phrases and the associated adjectives possibly preceded with adverbs. This consists of the following process.

5.3.4.1 Information Component Extraction

The information component extraction mechanism is implemented as a rule-based system which analyzes dependency tree to extract information components.

5.3.4.2 Feature and Opinion Extraction

The noun and adjective phrases are eliminated because of the design of information component and it is found to be necessary to consolidate the final list of information components. The consolidation process consists of two stages. In the first stage a feasible collection of product features is identified using term frequency (tf) and inverse document frequency (idf). In the second stage for each product feature the list of all opinions and modifiers is compiled that are used later for polarity determination of the opinion sentences. The tf-idf value for each noun phrase is calculated using the equations (2) and (3).

$$tf - idf(t_i) = tf(t_i) \times idf(t_i) \tag{2}$$

$$idf(t_i) = \log\left(\frac{|D|}{|\{d_j : t_i \in d_j\}|}\right) \tag{3}$$

Where

tf(ti) = number of documents containing term ti

|D| = total number of documents

|{dj : ti € dj}| = number of documents where term ti appears.

5.3.5 Feature Visualizer and Ranker

Feature Visualizer and Ranker module involves the following process steps. In the first step the polarity of the extracted opinions was identified and they were classified as objective, positive and negative. The polarity is classified using Senti-Word Net, a lexical resource. In the second step the opinion sentences are examined for each feature. Based on the score values they are mapped to the positive or negative class. The objective class is not considered mostly because the users were interested in either positive or negative views rather than objective views. The max function is applied to decide the class of an opinion sentence. For all the features with their corresponding positive and negative opinions are maintained in a table. The polarity value of the opinion word is multiplied with the number of sentences which contain that opinion in order to obtain the overall weight of the feature as shown in equation (4).

$$Total\,Weight = \sum_{n-1}^{d}\begin{pmatrix} Weight\ of\ the\ positive\ features - \\ Weight\ of\ negative\ features \end{pmatrix} \tag{4}$$

where d is the number of documents which contain this feature along with a commented word.

6. RESULTS AND DISCUSSION

In this section we have provided the results of our proposed methodology and have analyzed their performance. A Java program is written to extract the positive and negative features from the hospital feedback database.

The positive and negative opinions of the features were extracted from the database. Some of the positive and negative opinions extracted and their scores calculated were shown in Table 1 and Table 2.

Table 1: Tabular column for Positive opinions

Features	Positive Opinion	Modifier	Score
Patient Care	Excellent		5
Patient Care	Good	Very	4.5
Medical Knowledge	Average	-	4
Medical Knowledge	Agreeable		4
Professionalism	Exceptional		5
Professionalism	Incomparable		5
Timely service	Fine		2
Timely service	Normal		2

Table 2: Tabular column for Negative opinions

Features	Negative Opinion	Modifier	Score
Patient Care	Bad	Very	-0.5
Patient Care	Unsatisfied	-	-1
Medical Knowledge	Nothing	-	-2
Medical Knowledge	Disgusting		-2
Professionalism	Unconvinced	-	-1
Timely service	Disappointed		-1
Timely service	Worst		-2

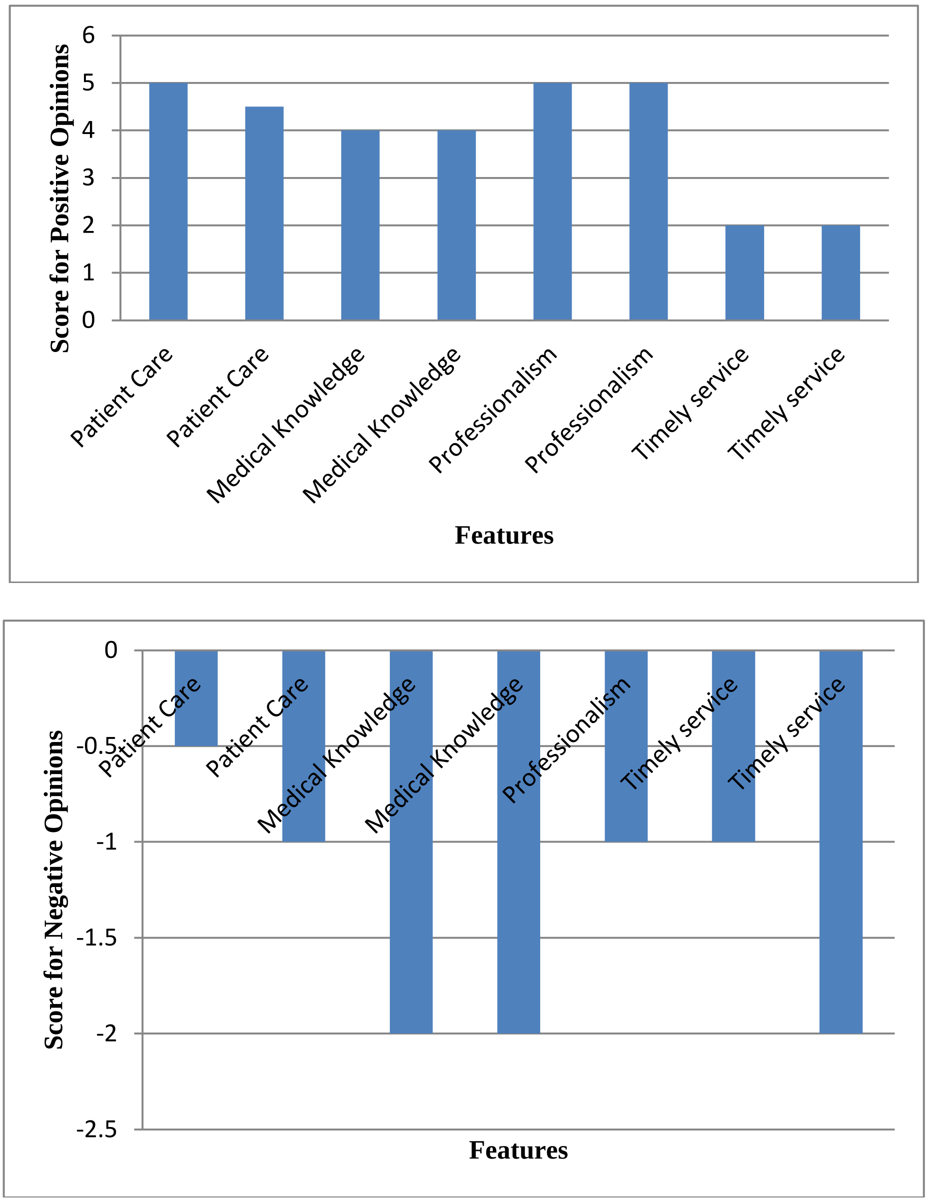

Fig. 3: Positive and negative opinion

The graphical plot for the positive and negative opinions respectively in the X-axis and their scores in the Y-axis was given in Fig..3. The result obtained from our proposed method is given in Fig. 4.

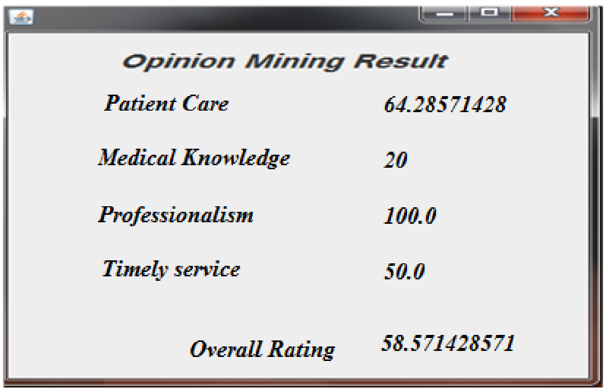

Fig. 4: Output obtained from the Proposed Methodology

7. CONCLUSION

The expression of opinions of users in specialized sites for evaluation of products and services, and also on social networking platforms, has become one of the main ways of communication, due to spectacular development of web environment in recent years. The large amount of information on these platforms make them viable for use as data sources, in applications based on opinion mining and sentiment analysis. Our proposed methodology had 4 major phases such as 1) Data Pre-processing 2) Content Discovery 3) Opinion extraction 4) Opinion mining. Initially an unstructured data was extracted from the web document. This phase is used to format the data before sentiment analysis and mining. In the second phase two processes were carried out i.e., Feature extraction and opinion extraction. The features like term frequency, Part of Speech (POS) were extracted from the words in the documents. After feature extraction, the useful information related to the item's features and use it to rate them as positive, neutral, or negative were extracted. This final phase was done by supervised learning algorithm decision tree classifier with the help of features extracted. In the final step ranking and classification was done. As an extension of the research presented in this paper we want to improve the algorithm, enriching the training set of examples, on the way, with examples classified as strong positive or negative, by an established score of classification.

8. REFERENCES

[1] Jack G. Conrad, Frank Schilder, 2007 , “Opinion mining in legal blogs”,In Proceedings of the 11th International Conference on Artificial Intelligence and Law (ICAIL07).

[2] KushalDave ,SteveLawrence,David M. Pennock, 1995, “Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews”.

[3] Anshu Jain, Dr. Suresh Jain,Mrs. PragyaShukla ,Hitesh Bandiya,, 2012, “Towards Automatic Detection Of Sentiments In Customer Reviews”, International Journal of Information Sciences and Techniques (IJIST) Vol.2, No.4.

[4] Liu, B., Hu, M., and Cheng, 2005, “Analyzing and Comparing Opinions on the Web.”, In Proceedings of the 14th International World Wide Web conference.

[5] Pang, B., and Lee, 2008, “Opinion Mining and Sentiment Analysis”, Foundations and Trends in Information Retrieval.

[6] Hu, M., and Liu, 2004, “Mining Opinion Features in Customer Reviews.”, In Proceedings of the 19th National Conf. on Artificial Intelligence.

[7] Miao, Q., Li, Q., and Dai, , 2008, “A Unified Framework for Opinion Retrieval”, In Proceedings of the Web Intelligence.

[8] Miao, Q., Li, Q., and Dai, 2008, “An Integration Strategy for Mining Product Features and Opinions.”’ In Proceedings of the 2008 ACM CIKM International Conference on Information and Knowledge Management.

[9] Archak, N., Ghose, A., and Ipeirotis, 2007, “Show me the money!: deriving the pricing power of product features by mining consumer reviews”, In Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

[10] TanvirAhmad , Mohammad NajmudDoja, 2012, “Ranking System for Opinion Mining of Features from Review Documents”, IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 1.

[11] Qingliang Miao, Qiudan Li, Daniel Zeng, 2010, “Mining Fine Grained Opinions by Using Probabilistic Models and Domain Knowledge”, ACM International Conference on Web Intelligence and Intelligent Agent Technology IEEE.

[12] Ion Smeureanu, CristianBucur, 2012, “Applying Supervised Opinion Mining Techniques on Online User Reviews”, InformaticaEconomică vol. 16, no. 2.

[13] Hsinchun Chen, 2010, “AI and Opinion Mining, Part 2”, IEEE.

[14] G.Vinodhini, RM.Chandrasekaran, 2012, “International Journal of Advanced Research in Computer Science and Software Engineering”, Volume 2, Issue 6.

[15] Qi Su1, XinyingXu, HongleiGuo, ZhiliGuo, Xian Wu, Xiaoxun Zhang, BinSwen and Zhong Su, 2008, “Hidden Sentiment Association in Chinese Web Opinion Mining”, IW3C2.

[16] LI Caiqiang and Ma Junming, 2012, “Research on Online Education Teacher Evaluation Model based on Opinion Mining”, National Conference on Information Technology and Computer Science.

[17] S. Wang and H. Wang, 2008, "A Knowledge Management Approach to Data Mining", Management and Data Systems, Vol. 108, No. 5, pp. 622-634.

[18] B. Dhanalakshmi, A. Chandra Sekar, 2013, “Fast and Efficient Opinion Mining Technique for Online Reviews using Ranking and Classification Algorithms”, International Review on Computers and Software, Vol.8, No.7, pp. 1587-1595.