Distributed Sentiment Embedding and Emoticon Based Opinion Mining

K. Varada Raj Kumar
Assistant Professor
D. Divya Manasa
M.Tech Student
B. S. B. P. Rani
Assistant Professor
SIR C R Reddy College of Engineering, Eluru, West Godavari Dt, AP, India
.

Abstract-Despite the rapid growth in social network sites and in data mining for emotion (sentiment analysis), little research has tied the two together and none has had social science goals. This article examines the extent to which emotion is present in MySpace comments, using a combination of data mining and content analysis, and exploring age and gender. A random sample of 819 public comments to or from U.S. users was manually classified for strength of positive and negative emotion. Two thirds of the comments expressed positive emotion but a minority (20%) contained negative emotion, confirming that MySpace is an extraordinarily emotion-rich environment. Females are likely to give and receive more positive comments than males, but there is no difference for negative comments. It is thus possible that females are more successful social network site users partly because of their greater ability to textually harness positive affect.

Keywords- Opinion Mining, Emotion, Sentiment Analysis, Prediction techniques

I. INTRODUCTION:

The computer-aided detection, analysis and application of emotion, particularly in text, has been a growth area in recent years (Pang & Lee, 2008). Almost all of this research has focused on detecting opinions in large bodies of text. For example, a program might scan a large number of customer comments or reviews of a manufacturer's products and report which aspects of which products tended to receive positive and negative feedback. Known as opinion mining (computer science) or sentiment analysis (computational linguistics), this approach typically works by identifying positive words or phrases in free text (e.g., "I like", or "rocked!") and tying them to the objects referred to (e.g., "the leather seats", "the package of extras"). From a wider social perspective, emotion is important to human communication and life and so it seems that the time is ripe to exploit advances and intuitions from opinion mining in order to detect emotion in a wider variety of contexts and for primarily social rather than commercial goals. In particular, is it now possible to detect emotion in people's textual communications and use this to gain deeper insights into issues for which emotion can play a role? For instance, how important is emotional expression for: effective communication between friends or acquaintances, winning an online argument, automatically detecting abusive communication patterns in chat rooms, or detecting predatory behavior online?

This article begins the process of moving from opinion mining to emotion detection by using a case study of MySpace comments to demonstrate that it is possible to extract emotion-bearing comments on a large scale, to gain preliminary results about the social role of emotion and to identify key problems for the task of identifying emotion in informal textual communications online. Hence, although it is preliminary and exploratory it is designed to report useful information for future emotion detection research and for those interested in social network communication. Large scale data collection and analysis from social network sites has already been used for social science research goals (Kleinberg, 2008) but not yet in combination with emotion detection.

Company strategies, marketing campaigns, and product preferences. Many new and exciting social, geo political, and business-related research questions can be answered by analyzing the thousands, even millions, of comments and responses expressed in various blogs (such as the blogosphere), forums (such as Yahoo Forums), social media and social network sites (including YouTube, Face book, and Flikr), virtual Worlds (such as Second Life), and tweets (Twitter). Opinion mining, a sub discipline within data mining and computational linguistics, refers to the computational techniques for extracting, classifying, understanding, and assessing the opinions expressed in various online news sources, social media comments, and other user-generated content. Sentiment analysis is often used in opinion mining to identify sentiment, affect, subjectivity, and other emotional states in online text.

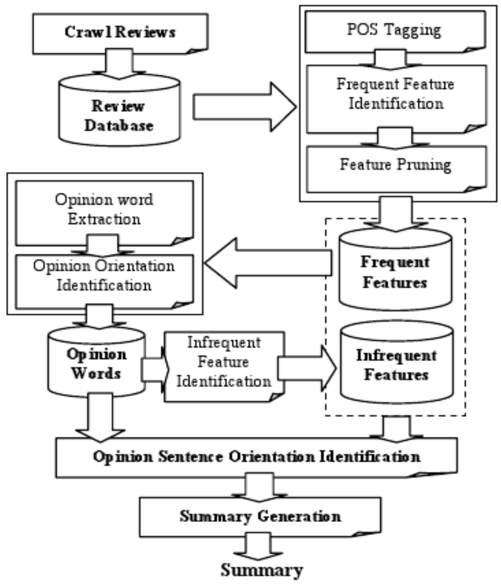
Our work builds on previous studies focusing on the relationship between the discussions held in firm-specific finance Web forums and public stock behavior. However, instead of assuming a shareholder view of participants in a finance Web forum as in previous research, and considering them to be uniformly representative of investors, we adopted a stakeholder perspective. This perspective more accurately represents the diversity of the constituency groups participating in the Web forum and closely aligns the analysis with the corporation's stakeholder theory. To address the broad questions posed in this research, and guided by the literature reviewed, we developed a framework for analysis with four major stages: stakeholder analysis, topical analysis, sentiment analysis, and stock modeling. During the stakeholder analysis stage, we identified the stakeholder groups participating in Web forum discussions. In the topical analysis stage, the major topics of discussion driving communication in the Web forum are determined. The sentiment analysis stage consists of assessing the opinions expressed by the Web forum participants in their discussions. Finally, in the stock modeling stage, we examine the relationships between various attributes of Web forum discussions and the firm's stock behaviour.

Our work builds on previous studies focusing on the relationship between the discussions held in firm-specific finance Web forums and public stock behavior. However, instead of assuming a shareholder view of participants in a finance Web forum as in previous research, and considering them to be uniformly representative of investors, we adopted a stakeholder perspective. This perspective more accurately represents the diversity of the constituency groups participating in the Web forum and closely aligns the analysis with the corporation's stakeholder theory.

To address the broad questions posed in this research, and guided by the literature reviewed, we developed a framework for analysis with four major stages: stakeholder analysis, topical analysis, sentiment analysis, and stock modeling. During the stakeholder analysis stage, we identified the stakeholder groups participating in Web forum discussions. In the topical analysis stage, the major topics of discussion driving communication in the Web forum are determined. The sentiment analysis stage consists of assessing the opinions expressed by the Web forum participants in their discussions. Finally, in the stock modeling stage, we examine the relationships between various attributes of Web forum discussions and the firm's stock behavior.

According to the processing grain, opinion mining could be divided into three levels: document level, sentence level and feature level. For the opinion mining on document and sentence level, the task is to classify either positively or negatively in a review. However, the sentiment orientation of a review is not sufficient for many applications. Opinion mining begins to focus on the finer-grained features level mining. The task is to find not only the sentiment orientation but also the commented features. This information could be used to deeply analyze prevalent attitudes or generate various types of opinion summaries. This paper focuses on the feature level product reviews mining. Given a review, the task is to extract product feature associated with its sentiment orientation. The task is typically divided into three main subtasks: identifying product features, identifying opinions regarding the product features, and determining the sentiment orientation of the opinions.

II. ARCITECHTURE



The product features are mostly noun or noun phrases, so we regard this subtask as an entity recognizing process, and hope to transfer the effective NER techniques to solve this problem. We adopt the Conditional Random Fields module (Lafferty et al, 2001) to implement this subtask, which has been proved well performance in information extraction field. CRFs modules has the advantages of relaxing strong independence assumptions made in HMM and avoiding the label bias problem existed in MEMM. In the CRFs modules, we import word, POS and semantic information as tokens. The semantic information not only includes its character as a product feature, but also contains the character about opinion expression. The opinion information is a good indicator, because people like to express their opinions around the product features. All the semantic information is captured dependent on the above language resources. In this stage, we not only tag the product features but also tag the opinion words as attachment. Another reason for us to adopt the supervised method to implement this subtask is that the unsupervised frequency-based methods are dependent on the statistic of the corpus, when given a single sentence, they couldn't execute effectively.

Posting opinions:

In this module, we get the opinions from various people about business, e-commerce and products through online. The opinions may be of two types. Direct opinion and comparative opinion. Direct opinion is to post a comment about the components and attributes of products directly. Comparative opinion is to post a comment based on comparison of two or more products. The comments may be positive or negative.

Object identification:

In general, people can express opinions on any target entity like products, services, individuals, organizations, or events. In this project, the term object is used to denote the target entity that has been commented on. For each comment, we have to identify an object. Based on objects, we have to integrate and generate ratings for opinions.

The object is represented as “O”. An opinionated document contains opinion on set of objects as $\{o_1, o_2, o_3 \ldots o_r\}$.

Feature extraction:

An object can have a set of components (or parts) and a set of attributes (or properties) which we collectively call the features of the object. For example, a cellular phone is an object. It has a set of components (such as battery and screen) and a set of attributes (such as voice quality and size), which are all called *features* (or *aspects*). An opinion can be expressed on any feature of the object and also on the object itself.

Opinion-orientation determination:

The opinion holder is the person or organization that expresses the opinion. In the case of product reviews and blogs, opinion holders are usually the authors of the posts.An opinionon a feature *f (*or object *o*) is a positive or negative view or appraisal on *f* (or *o*) from an opinion holder. Positive and negative are called opinion orientations. From this opinion orientation we have to determine the type of opinion whether it is direct opinion or comparative opinion.

Direct opinion:

A *direct opinion* is a quintuple $(o_j, f_{jk}, oo_{ijkl}, h_i, t_l)$, where$o_j$is an object, $f_{jk}$is a feature of the object o_j, $oo_{ijkl}$is the orientation of the opinion on feature $f_{jk}$of object o_j, $h_i$is the opinion holder, and $t_l$is the time when the opinion is expressed by h_i. The opinion orientation oo_{ijkl} can be positive, negative, or neutral.

Comparative opinion*:* A*comparative opinion* expresses a preference relation of two or moreobjects based their shared features. A comparative opinion is usually conveyed using the comparative or superlative form of an adjective or adverb, such as “Coke tastes better than Pepsi.”

In this subtask, nearest vicinity match based and syntactic tree based methods are proposed to confirm the associated opinion word. As observed, the opinion words mostly appear around the features in the review sentences. They are highly dependent on each other. So we hypothesize that opinion words appear around product features. If an opinion word co-occurs with a product feature within a given distance in a sentence, this opinion word is regarded to be associated with this product feature. Otherwise they are considered to be unrelated. Nearest vicinity match based method has two steps to identify the opinion words. First, it takes the product feature as the center to find opinion word tagged by CRFs in the given distance. If there is no opinion word tagged by CRFs, then it secondly looks at the opinion lexicon for the further search. If there is also no opinion word found, the product feature is considered to have no sentimental meaning, which will be deleted. Dependent on the plane distance to capture the opinion words is not

With these concepts in mind, we can define an object model, a model of an opinionated text, and the mining objective, which are collectively called the *feature-based sentiment analysis model*. In the *object model*, an object “O” is represented with a finite set of features,

$F = \{f_1, f_2, \ldots, f_n\}$

which includes the object itself as a special feature. Each

feature $f_i \in F$ can be expressed with any one of a finite set of words or phrases $W_i = \{w_{i1}, w_{i2}, \ldots, w_{im}\}$

Which are the feature’s *synonyms*.

sufficient. So we adopt syntactic tree based method to capture the relation. Here, we compute the distance of two items based on the syntactic parsing tree, and measure it by the shortest path. Figure 1 shows the example of a review. It could not be determined by nearest vicinity match based method which opinion word is associated with the feature. As in the sentence, word has the same plane distance with both the opinion word and opinion word. Nearest vicinity match based method is dependent on the distance of the text string to judge the relative extent of two terms. It has no consideration for grammar information of the sentence. In fact, the grammar and syntactic structure contains more associated information between the terms. So we utilize the distance of the two terms in the parsing tree to measure their relation. The distance of two leaf nodes are calculated with the shortest path of the two nodes in parsing tree. The distance of word with opinion word is 7, and that with opinion word is 9. The opinion word is more associated with feature

III. BACKGROUND

Usenet has been providing a means for people across the world to participate in online discussions on variety of topics since the early 1980’s, and still continues to draw thousands of postings a day. With so many postings and a great number of different threads proceeding concurrently, it is a great challenge to identify the set of postings or threads that relate to a specific topic one is interested in. Current Usenet clients can group postings by the thread they belong to, but they do not provide any further organization. . The important technical details discussed in newsgroups are much desired, and give better answers particularly to technical questions than searching the web does. However it is a daunting task to search the complete list of newsgroups, many of which may be talking about the same topic, and browse through all the threads searching for the right ones.

A system that could automatically process the newsgroup and generate a list of topics being talked about could be useful to give a brief synopsis of the whole newsgroup and also for advanced search. Even a goal based traditional IR system could be significantly improved, if backed by some classification or metadata about the newsgroup postings. Also, given the weights of the topics and their relative similarity measures among each other as well as between documents, one can generate a visualization of the newsgroup threads and their topics, providing a way to both visualize and navigate the clustered threads easily.

We implemented a system to automatically extract newsgroup threads and categorize them based on most prominent topical categories. We implement and compare two methods for feature extraction from newsgroup threads, one based on purely statistical methods and the other making use of semantic information in the newsgroup threads in order to build a topical structure for a newsgroup corpus. We compared these two methods, viz. chi-square based feature extraction and Latent Semantic Analysis, to determine which method will yield better results. Our initial hypothesis was that the Latent Semantic Analysis will be able to extract more sensible topics, as suggested by many of the literatures, but in practice the difference in results from the two methods was not very significant.

The similarity measure used is a simple vector space cosine measure, and the similarity parameter can be tweaked to get the desired grouping accuracy. The document vectors obtained as a result of the topic extraction step are *m*-dimensional vectors, where *m* is the total number of term in the newsgroup. The topic clustering step calculates the document-document similarity measure matrix using cosine similarity between document vectors, and generates a sorted list of document pairs based on their similarity. We then find the 2 most similar topics and merge then using proper weights, taking into account the number of previous merges, so as to generate the correct mean value of the cluster at each step. This process is repeated until the maximum similarity measure amongst all remaining document vectors becomes less than the predefined threshold value. Another way to break out of the clustering could be after we get the desired number of clusters. Right now the similarity value is set at 10%, because of the sparseness of data .This section reviews several aspects of the background to automatic emotion detection in social network sites: opinion mining (i.e. automatic opinion detection); the psychology and sociology of emotion (because emotion is a complex construct); and social network communication and usage. Gender differences in emotion and language are also discussed.

Opinion Mining and Text Mining

Opinion mining or sentiment analysis is the automatic detection of opinions from free text. This research area has been partly motivated by the commercial goal of giving cheap, detailed and timely customer feedback to businesses (Pang & Lee, 2008). Before the Internet, businesses would have to rely upon relatively slow and expensive methods of gaining customer feedback, such as phone or mail surveys, interviews and focus groups. Online, however, they may be able to gain feedback from online customer reviews, blogs, comments and chatroom discussion, assuming that a computer program can filter out the relevant data from the rest of the web or a particular reviews website. In this context, the goal of opinion mining is to identify positive and negative opinions in free text and to associate this opinion with relevant objects. The goal might be detail in the sense of identifying what is discussed and how (e.g., which aspects of a car are liked or disliked), or the goal might be a judgment in the sense of diagnosing the nature and strength of opinion (e.g., diagnosing how much a reviewer liked a film from their online review).With the growing interest in opinion mining from web data, more works are focused on mining in English and Chinese reviews. Probing into the problem of product opinion mining, this paper describes the details of our language resources, and imports them into the task of extracting product feature and sentiment task. Different from the traditional unsupervised methods, a supervised method is utilized to identify product features, combining the domain knowledge and lexical information. Nearest vicinity match and syntactic tree based methods are proposed to identify the opinions regarding the product features. Multi-level analysis module is proposed to determine the sentiment orientation of the opinions. With the experiments on the electronic reviews of COAE 2008, the validities of the product features identified by CRFs and the two opinion words identified methods are testified and compared. The results show the resource is well utilized in this task and our proposed method is valid.

Opinion mining is often split into two consecutive tasks: detecting which text segments (e.g., sentences) contain opinions and the polarity and perhaps strength of that opinion (Pang & Lee, 2008). A simple technique counts how often positive and negative words occur or how often they co-occur in sentences with given target terms (e.g., “engine reliability”). Whilst full machine comprehension of text is currently impossible, computational linguistics techniques can partly analyse the structure of text, using it to more accurately detect sentiment. This approach might incorporate negating words (Das & Chen, 2001) like “not”, booster words like “very” and grammatical structures common in sentiment-bearing sentences (Turney, 2002). It relies upon reasonably grammatically correct English to function effectively, however, which makes it less useful in environments like social network sites with much informal language. Many refinements of the above approaches have been proposed (e.g., Konig & Brill, 2006; Turney, 2002).

Text mining applications have also been developed in psychology, communication studies, management and corpus linguistics (for a review see: Pennebaker, Mehl, & Niederhoffer, 2003). For instance, some psychological disorders can be quite reliably diagnosed in patients based upon a simple word frequency analysis of speech (Oxman, Rosenberg, & Tucker, 1982); political statements (Hart, 2001) and business mission statements (Short & Palmer, 2008) have been analysed for the strength of variables including optimism; and a factor analysis across a wide range of text genres has identified that the degree of author involvement in a text as opposed to an informational orientation (arguably a weak expression of emotion) is something that tends to be constant within genres but varies between genres (Biber, 2003).

The psychology of emotion

Many different aspects of emotion can be measured, including: individuals' self-reports of feelings, neurological changes, autonomic system reactions, and bodily actions – including facial movements (Mauss & Robinson, 2009). These seem to overlap to between different emotions leading them to be described as syndromes rather than clear sets of identifiable features. Eckman (1992) and others nevertheless argue that there are basic or fundamental emotions that are relatively universally recognised and apparently experienced by humans, and that these exist as a result of evolutionary pressure. For example, autonomic changes and cognitive processes during fear prepare a person to run away from danger. In support of this, there is scientific evidence that at least five different emotions (fear, disgust, anger, happiness, sadness) are demonstrably different in the sense of activating different combinations of brain regions (Murphy, Nimmo-Smith, & Lawrence, 2003); adding surprise gives Ekman's (1992) main list of six basic emotions. Eckman's (1992) evidence found in support of emotions being basic is a set of six general characteristics common to all basic emotions (e.g., brief duration, presence in other primates) and three types of characteristic that exist but differ between emotions: signals (e.g., facial expressions); physiology (e.g.,autonomic nervous system activity patterns); and antecedent events (e.g., a dangerous event occurring).

Weight Mechanism for Words:

The weight for $a_x = (a_{x1}, \dots a_{xd})$ is

$w(a_x) = wt(a_x) + ws(a_x)$

where $wt(a_x)$ is the weight for feature a_x in the training data, given that v and w are part of the set of c class labels, $v \neq w$, and $c \geq 2$:

$$wt(a_x) = \max_{v,w}\left(P(a_x \mid v) \log\left(\frac{P(a_x \mid v)}{P(a_x \mid w)}\right)\right)$$

and $ws(a_x)$ is the semantic weight for feature a_x:

$$ws(a_x) = \frac{1}{d}\sum_{i=1}^{d}\left(\frac{1}{k}\sum_{j=1}^{k} s(a_{xi}, j)\right)$$

where $s(a_{xi}, j)$ is the sum of the positive and negative scores for the word a_{xi} and j is one of the k senses of a_{xi} in SentiWordNet.

The above list excludes some emotions considered important by others, such as anxiety, guilt, shame, envy, jealousy, compassion and love (Lazarus, 1991, p. 122). Non-basic emotions are sometimes seen as combinations of basic emotions and seem to vary more between cultures. Emotion perception is culture-specific because some societies describe emotions never apparently experienced elsewhere (e.g., the oft-mentioned "state of being a wild pig" (Newman, 1964) in a New Guinea community).

From the perspective of felt human experiences rather than at the neurological or descriptive levels, it seems that there are *two* fundamental dimensions rather than a range of differing kinds of emotions (Fox, 2008, p. 120). First, the *valence* of an experienced emotion is the degree to which it is strongly positive or negative. Second, the level of *arousal* felt is the amount of energy perceived (e.g., from lethargic to hyperactive). This assertion apparently contradicts the neurological evidence above of at least five emotions and the linguistic evidence in the form of the existence of a wide range of non-synonymous terms for emotions. Nevertheless, research has shown that people describing the same traumatic event may use a wide range of different emotional terms (e.g., sad, angry, upset) almost indiscriminately (Barrett, 2006) and that the two dimensions of valence and arousal seem to be the key underlying factors. A consequence of this is that identifying valence and arousal is likely to be far easier and more reliable than other types of emotion detection.

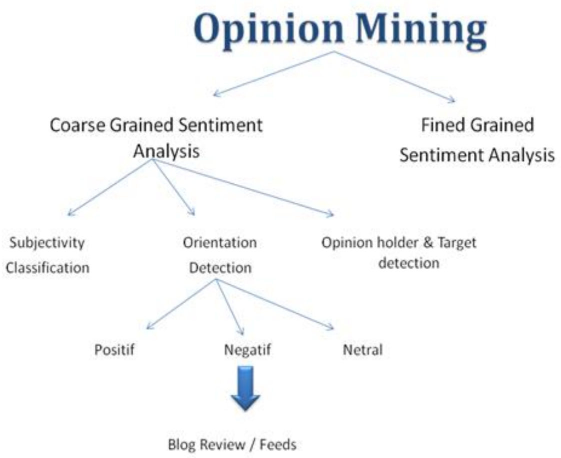


Fig1

A Tree structured opinion mining classification techniques to obtain efficiency over online social networks.

Almost contracting the valence-arousal model of emotion perception, there is evidence that levels of positive and negative emotion are not correlated: a person can simultaneously experience varying levels of both, although they may be perceived as separate simultaneous emotions (Watson, 1988) – for example, enjoying the fear in bungee-jumping or missing a loved one.

Importantly for emotion classification in the current paper, individuals perceive and react to potentially emotional stimuli in significantly different ways. Personality differences impact the strength of emotion perceived from a stimulus and the tendency to perceive a negative or positive context when there is a choice. The latter broadly reflects a pessimistic or an optimistic person. More specifically, two of the five commonly recognised personality traits in psychology are associated with the ability to experience emotion: extraversion with positive emotions and neuroticism with negative emotions (Fox, 2008, p. 53-58). It has also been shown that people react in different ways even to clear emotion expression devices, such as emoticons, in the sense of drawing inferences about the characters of the users (Fullwood & Martino, 2007). In consequence, irrespective of life experiences, people with different personalities are likely to disagree about the strength and polarity of emotion in many situations.

The sociology of emotion

In addition to psychological research and the social psychological perspective of identifying social and cultural factors in emotion expression, emotion has been extensively studied in sociology. Many theories have been developed to explain the role of emotion in various situations. One review grouped these theories mainly into the following broad types: dramaturgical and cultural, regarding emotion as a performance by individuals to an audience; ritual, regarding emotion as an important outcome of ritual processes, not only religious ceremonies but also standardised procedures used in human interaction; symbolic interactionist, regarding emotions as sometimes generated when individuals' self-identities are threatened or reinforced; symbolic interactionist with a psychoanalytic focus, analysing strategies used to deal with lack of confirmation of identity; exchange theories, regarding emotions either as commodities to be exchanged or as the outcome of exchanges; structural theories, based upon social power structures; and evolutionary theories, explaining current emotions on the basis of evolutionary social pressures (Turner & Stets, 2005, p. 23-25). Additional groups of theories have also been proposed (Stets & Turner, 2007). All of the above theories could add something to the understanding of emotion in communication between pairs of friends in social network sites, but they are mostly of peripheral relevance. For example, they typically involve groups of people rather than individuals or dyads. Whilst several theories are particularly promising on the surface, ritual theories tend to emphasise aspects of face-to-face communication, or co-presence, that are irrelevant here; dramaturgical theories are "best applied to situations where group solidarity is built up"; exchange theories tend to focus on emotion as the outcome of exchanges or networks of relationships (Turner & Stets, 2005, p. 99). Nevertheless, the approaches and ideas introduced in the emotion of sociology are important to emphasise that the expression of emotion is not just a simple description of a person's internal emotions but may be used by them in ways that are influenced by the people around them, their strategic goals, and previous experiences.

IV.CONCLUSION

The orientation of the product features are judged from multi-levels: sentence level, context level and opinion word level. Sentence level judgment considers whether a sentence express any opinion information. The features in non sentiment sentence should not be extracted. Context level judgment considers the sentiment transition by emotional adverbs or phrases, such as word "不(no)". Opinion word level judgment considers the opinion word associated with product feature. Since sentiment orientation is only positive or negative in this task, the results are combined of three level judgments with product. It is defined in the following way. Here, Is O(S) is a two value function, which judges the orientation of the feature by sentence level. If the sentence is judged to have no sentiment, then its value is 0, else is 1. At present, we only consider assumption sentence. We think this type of sentence doesn't express any opinion information, so the features in it should not be tagged. Is N (ai) is also a two value function, which considers the sentiment transition by emotional adverbs or phrases. It is on the context level to consider the orientation. If there is such a word around the feature in a region, the value of the function is -1, else is 1. Sign (oi) is directly the orientation of associated opinion word for the feature. Its value is defined as -1 for negative, 0 for neutral and 1for positive orientation. Both the information about ai and oi are determined by looking at the opinion words lexicon and factor words lexicon.Classification of opinion using advanced methodologies to identify and prediction of product rating, web ranking implementing comparative study and proof of concept that implemented techniques are more appropriate and efficient in sentimental analysis for best ranking techniques.

V. REFERENCES

[1]. E. Riloff, J. Wiebe, and T. Wilson," Learning Subjective Nouns using ExtractionPattern Bootstrapping," Proc.7th Conf. Natural Language Learning,ACM Press, 2003, pp. 25–32.

[2]. A. Abbasi et al., "Affect Analysisof Web Forums and Blogs using CorrelationEnsembles," IEEE Trans.Knowledge and Data Eng., vol. 20,no. 9, 2008, pp. 1168–1180.

[3]. E. Riloff, S. Patwardhan, and J. Wiebe,"Feature Subsumption for OpinionAnalysis," Proc. Conf. Empirical Methodsin Natural Language Processing, ACM Press, 2006, pp. 440–448.

[4]. Y. Dang, Y. Zhang, and H. Chen,"A Lexicon-Enhanced Method for SentimentClassification: An Experimenton Online Product Reviews," IEEEIntelligent Systems, vol. 25, no. 4,pp. 46–53.

[5]. T. Mullen and N. Collier, "Sentiment Analysis Using Support Vector Machines with Diverse Information Sources," Proc. Conf. Empirical Methods in Natural Language Processing,ACM Press, 2004, pp. 447–454.

[6]. Ankit Singh, and Md. Enayat Ullah, "Aspect based Sentiment Analysis".

[7]. Maria Giatsoglou , Manolis G. Vozalis , Konstantinos Diamantaras , Athena Vakali , George Sarigiannidis , Konstantinos Ch. Chatzisavvas , " Sentiment analysis

leveraging emotions and word embeddings" , Expert Systems With Applications (2017).

[8]. Minqing Hu and Bing Liu ," Mining Opinion Features in Customer Reviews and Mining and Summarizing Customer Reviews".

[9]. Bing Liu, "Opinion Mining by Using Multinomial Naive Bayes on Blog","Sentiment Analysis: mining sentiments, opinions, and emotions".