

Study of Automatic Extraction, Classification, and Ranking of Product Aspects Based on Sentiment Analysis of Reviews

Muhammad Rafi, M.Rafay Farooq, Usama Noman, Abdul Rehman Farooq, Umair Ali Khatri

Abstract--- It is very common for a customer to read reviews about the product before making a final decision to buy it. Customers are always eager to get the best and the most objective information about the product they wish to purchase and reviews are the major source to obtain this information. Although reviews are easily accessible from the web, but since most of them carry ambiguous opinion and different structure, it is often very difficult for a customer to filter the information he actually needs. This paper suggests a framework, which provides a single user interface solution to this problem based on sentiment analysis of reviews. First, it extracts all the reviews from different websites carrying varying structure, and gathers information about relevant aspects of that product. Next, it does sentiment analysis around those aspects and gives them sentiment scores. Finally, it ranks all extracted aspects and clusters them into positive and negative class. The final output is a graphical visualization of all positive and negative aspects, which provide the customer easy, comparable, and visual information about the important aspects of the product. The experimental results on five different products carrying 5000 reviews show 78% accuracy. Moreover, the paper also explained the effect of Negation, Valence Shifter, and Diminisher with sentiment lexicon on sentiment analysis, and concluded that they all are independent of the case problem, and have no effect on the accuracy of sentiment analysis.

Index Terms—Aspect ranking, Product Aspect Ranking, Sentiment analysis, Sentiment lexicon

I. INTRODUCTION

Web 2.0 is rich on user-generated contents for different products. For example, CNet.com involves more than seven million product reviews, whereas Pricegrabber.com contains millions of reviews on more than 32 million products. Consumers nowadays believe to buy more products from online store than any physical store. Ever since the number of customers who prefer to shop online increases, the phenomenon of asking

reviews about the product from family and friend also increased. Now with the advancement of technology, the task of asking for reviews from friends and family has shifted to Product Reviews website. Seventy percent consumers believe that the online reviews are the most trusted and reliable source of information [1]. Although web has a large collection of information, sifting through these texts and extracting valuable information from these disorganized reviews is very challenging and daunting task, but can be solved using sentiment analysis.

Sentiment analysis carries great importance in digital world. Identifying sentiments from the natural text is not very difficult but tricky. Correct analysis of reviews could increase the sale of a company to 200% in a month; therefore, many researchers are experimenting with different methods to find the complete solution of this daunting task. The main challenge in finding sentiments from the text is to find its scope and its intensity. Since mostly reviews are subjective and carry ambiguous opinion, it is very hard to find opinion words, understand their contextual meaning, and identify their scope. Due to non-uniform structure of web, it is also very challenging to mine the information the reader actually needs. Online reviews constitute a small part of this huge clustered web pages.

II. LITERATURE REVIEW

The rapid growth in e-commerce is due to increasing trust of online customers. There are millions of products, from thousands of manufactures and distributors available for sale online. Every category has hundreds of product to choose from, and it is very difficult for an online buyer to make a wise decision. Therefore, buyers go through reviews about the product to make a final decision, but due to subjectivity and ambiguity of reviews, it often does not reach any valuable conclusion. The difficult part of this activity is to search this distributed information from multiple website; analyze the subjectivity of text, and conclude the final notions about the product from these reviews.

Recent study [2] showed the impact of reviews on the sale of product, and opinion of the customer. In this work, a special summarization technique is presented which was different from traditional text summarization because it focuses only on important aspects of the product rather than summarization of the whole review. This summary proved to be very beneficial for the online users who are about to make a purchase. Another recent work is of Zheng-Jun Zha, Jianxing Yu, Meng Wang, and Tat-Seng Chua [3] related to the identification of product aspects. The work proposed a framework that can rank the important aspects of the product by exploiting product's aspects frequency and its probability in a review. Their algorithm showed significant

Manuscript received June 30, 2015; revised August 08, 2015. The Computer Science Department of National University of Computer and Emerging Sciences supported this work.

Muhammad Rafi is teaching as an Assistant Professor of CS Department on National University of Computer and Emerging Sciences, ST-4 Sector 17-D, On National Highway, Karachi. (UAN: (021) 111 128 128; Tel: (92-213) 4100541-6; Fax: (92-213) 4100549; email: Muhammad.rafi@nu.edu.pk)

Muhammad Rafay Farooq is the Software Engineer at Folio3 Software House, Karachi, Pakistan (e-mail: k112257@nu.edu.pk).

Usama Noman is currently incubate at NEST I/O and working on start-up project in Karachi, Pakistan (email: k112119@nu.edu.pk).

Abdul Rehman Farooq is the Software Engineer at Techlogix Software House, Karachi Pakistan (email: k112017@nu.edu.pk).

Umair Ali Khatri is currently incubate at NEST I/O and working on start-up project in Karachi, Pakistan (email: k112224@nu.edu.pk).

improvements over other already proposed methods when tested on 95 thousands customer reviews. Identifying the sentiments associated with different aspect of product seems to be a very instrumental in overall sentiment of the product.

Some researchers developed their own sentiment lexicons, and devised tools that could automatically extract reviews, and find important features based on supervised learning techniques. Bo Pang, Lillian Lee, and Shivakumar were among the very few early researchers who proposed the technique of sentiment classification using machine learning [4]. They proposed a method, which can find the sentiment of a document not by topic but by overall sentiment i.e. if the review is positive or negative. They also highlighted the importance of unigram words in sentiment classification technique in their paper. Hanhoon Kang, SeongJoonYoo, and Dongil Han proposed a new lexicon for sentiment classification because of lack of sentiment words in already existing sentiment corpuses [5]. Their focus was to narrow the classification accuracy gap between positive and negative sentiment documents. They proposed a modified Naïve Bayes algorithm, which narrowed down the classification gap to 3.6% as compared with original Naïve Bayes. They chose the dataset of restaurant reviews for their experiment and concluded that unigrams and bi-grams features play a major role in sentiment analysis of reviews. Likewise, Kushal Dave, Steve Lawrence, and David M. Pennock developed an opinion-mining tool that can distinguish between positive and negative reviews automatically by assigning the features some scores based on heuristics [6]. The tool was prune to web based searches due to noise and ambiguity, and used supervised learning to find the important aspects from the text. In addition to that, Michael Wieg et.al presented a concrete summary in his paper on the role of negation in Sentiment Analysis [7]. They present computational approaches, and modeled the role of negation in sentiment analysis. In addition, they also discussed the limitation and challenges in negation modeling followed by the detection and scope of negative words. In comparison, some researchers discussed the impact of irony and sarcasm in online reviews and their effect on sentiment analysis task [8]. Few of them also explained the influence of negative and valence shifter words in determining the overall sentiment of a review. Elena Filatova presented a corpus generation experiment that can identify irony and sarcasm from a corpus in two levels: document level and text utterance where text utterance can range from a single sentence to a complete document [9]. Livia Polanyi and Annie Zaenen worked on valence shifters to determine the attitude of writers towards the material being described [10]. In contrast, Alistair Kennedy and Diana Inkpen use the help of contextual shifter to classify movie reviews [11]. They specifically examine three types of contextual valence shifters namely negations, intensifier, and diminisher, and studies their effect on classification of reviews. They did not assign weights to negative and positive words and treated all the words on the same level.

Our research mainly focused on how to identify important aspects of a product from its reviews, and rank those aspects based on their sentiment scores. Our work closely relates to the work of Zheng-Jun Zha [12], but instead of working on term frequency and probability, we used different lexicons to identify the score of aspect and later rank those aspects based on their sentiment scores. In this paper, we are

discussing a lexicon-based approach to find the sentiment score of aspects. The impact of sentiment lexicon on negation handling which is the most important part of sentiment analysis is also discussed. In summary, the main contribution of this research is as follows:

- i) We proposed a lexicon based approach to find the sentiment of product (aspect).
- ii) We analyzed the importance of a good opinion lexicon, and its effect on negation-handling task. We also concluded that if a good opinion lexicon is used, then we do not need to handle important features of linguistics like Valence shifter in the task of sentiment analysis.

Next section describes the proposed approach of finding the sentiment of aspects through lexicon. It describes the pre-processing task, challenges of extracting text from reviews, and the task of identifying relevant aspects and its impact on results. After that, the paper has the intermediate results of different lexicon with different window sizes. The final section describes negation handling, the effect of Valence shifter and Diminisher on sentiment scores and its effect on lexicon based approach.

III. THE PROPOSED APPROACH

In this section, we explained our proposed approach and working of the system. Before proposing our solution, we also presented a short summary of relevant challenges and problems faced previously. Our system has four phases namely Extraction of Reviews from Web, Identifying Aspects of Product from those reviews, Sentiment Analysis of Product (Aspect), and finally the ranking of product aspects based on sentiment score. In each phase, the output of one phase feed acts as an input for another. First, we extract reviews from the web and identify aspects by detecting noun phrases [13]. After finding aspects, we do sentiment analysis of words surrounded by those aspects, and assign score to each aspect. Finally, we rank all aspects based on their sentiment scores and present this information to the user. The following section will explain each phase of the system in detail.

A. Pre-Processing of Text

Preprocessing of text plays a vital role in the area of text classification and natural language processing. In order to get good results, this step plays a very important role in our system. The impact of pre-processing in the field of text classification is extensively studied, and research on various languages like Arabic, Turkish, and Portuguese [14], [15], [16] support our motivation behind doing pre-processing at this step. It has already proven that preprocessing takes almost 80% of the total time in classification process [17]. Many good techniques like TF/IDF, Stop word removal and stemming showed considerable impact on classification accuracy of documents with different domain dataset [18]. Experiments also conclude that different combination of preprocessing techniques should be applied instead of enabling or disabling them all to increase the accuracy [19]. In our approach, we removed stop words that expand sentiment word's domain and enhance discrimination degree between documents.

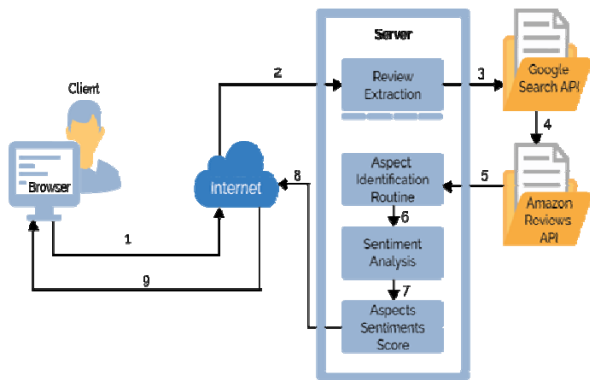


Fig 1. Proposed framework of the system

B. Extraction of Reviews from Web

The diversity of Web 2.0 makes the process of extracting relevant information from this unstructured and non-uniform spider of pages a very daunting task. For many natural language processing tasks, the size and quality of data used for training and testing is very crucial. Due to rapid increase of data around the web, it is very important to study how only the relevant data from different web sources can be extracted and processed for relevant tasks. Since most web pages contain tags and other non-content HTML characters [20], it is easy to extract content from any web page if we can exploit these tags. In our approach, we exploited HTML and XHTML tags and used python library for web scrapping to extract the reviews from different websites.

C. Aspect Identification

Identifying important and relevant aspects of the product is very important and the most sensitive part in sentiment analysis process. Proper identification of aspects is very challenging due to diversity of Natural Language Text. The most novel approach to identify important aspects from online reviews is to observe customer reviews and sort the frequency of aspects mentioned in most of the reviews. The majority class will represent the important aspect of that product [12]. For aspect identification, we find all the frequent nouns from the text [21] and sort them with their term frequency. For unigram and bi-gram aspects, we filter all Nouns with their term frequency. In addition, if we have two aspects with one having a common sub-string of other, we discard the shorter length aspect to be more specific. For example if we have two aspects “battery” and “battery life,” we will chose the second one because it is more specific than first. Table below shows some extracted unigram and bigram aspects of “iPod” by our system.

TABLE I
 UNIGRAMS OF IPOD

Battery	Songs	Music	Capacity	Software
Device	I-Tunes	Service	Computer	I-Pod
Controls	Interface	Audio	Bass	Adapter
Click Wheel	Sound Quality	Battery life	Storage capacity	Click Wheel
Battery replacement	Sleek design	Better technology	Color display	Battery replacement

D. Sentiment Analysis

Sentiment analysis and opinion mining are the most widely researched topics in the domain of Natural Language and Data mining. Different researchers have worked on different layers of sentiment analysis from Document level to sentence level and aspect level [22]. The important challenge faced in the area of sentiment analysis is the true meaning of opinion words. Likewise, it is not always necessary that an opinion word like “good” always carries some positive opinion when appears in a sentence. For example, the sentence “I want to buy a good camera, can you please give me some suggestion” does not carry any opinion about any camera but has opinion words. In our proposed methodology of finding the sentiment score of identified aspects, we used lexicon-based approach. We have experimented with different lexicons and tested different combinations to maximize the classification accuracy of sentiments. We used lexicons of **NRC Canada** [23] also known as Senti140, and **CSUIC** also known as Opinion Lexicon by Mingqing Hu et al. [13] in order to find the sentiment scores of the words surrounded by the aspect. The following section will explain the intermediate results and analysis when different lexicons are experimented on and tested.

Data Set

In our experiment, we have used the dataset of “Ming Lui spam detection in fake reviews” [24]. The dataset contains nine products with more than 5000 reviews in the whole collection. The reviews are manually annotated with ranking of each aspect identified. We train our algorithm in five products reviews, and find generalization accuracy. The format of the review is ASPECT [[+/-] RANK] ##REVIEW. Reviews were annotated and with each review, its aspect score was given. In addition to that, we also experiment with SemEval 2014 data set of Restaurant and Laptop reviews [25]. The dataset of each category is divided into two parts namely Training and Trialing. Laptop Training dataset contains 1900 reviews, whereas restaurant training dataset contains 1350 reviews all annotated by experienced annotators.

IV. EXPERIMENTS

In our approach, we used distance based linking to find how the opinion words affect the meaning of the target aspect. A windows size of four and opinion words from the target aspect on both sides are exhausted. We have applied distance based linking to aspect level and with each aspect; we stored the surrounding opinion phrase with window size of four. If there are more opinion words than aspects, then we apply the score of those opinion words to all the aspects coming in the window. For example, if the review is “The camera is very good and gives amazing and astonishing results.” There are three opinion words *very good*, *amazing*, *astonishing*, but only one aspect *camera*. Since all these opinion words come on the window size of four with camera, they will modify the sentiment score of camera. Following section will explain all the intermediate results of our experiment with different lexicons and window sizes.

First, we used only AFINN as sentiment lexicon and calculated sentiment scores of surrounding words in the distance of two. Our algorithm gave average accuracy of 26%. The major reason of this low accuracy was the short

length of this lexicon. There were only 2800 words in AFFIN and without handling negation and bi-words, the accuracy become worst. As a result, we add another rich lexicon named CSUIC aka Opinion Lexicon compiled by Bing Lui [24]. The results showed 6% improvement over AFFIN with overall accuracy to 32%. It is important to note that the window size of experiment was set to two during this analysis. After experimenting with windows size of two, we gradually increased our windows size from two to three and then to four. Major improvements are observed when window size was set to four. Mostly reviews contain long sentences due to which the window size of five decreased the accuracy. Moving further, we also experimented with different other lexicons like Sentiment 140 apart from CSUIC and AFFIN, and finally found the three best lexicons.

TABLE II
AFFIN, CSUIC, AND SENTI140 SCORE
WITH WINDOW SIZE = 3

Products	Distance 3	
	AFFIN+CSUIC+Senti140	CSUIC+Senti140
iPod	42.7%	45.8%
CanonG3	32.9%	34.3%
Hitachi Router	44.2%	47.2%
Norton	29.5%	30.3%
Average	37.7%	40.4%

Table III
AFFIN, CSUIC, AND SENTI140 SCORE
WITH WINDOW SIZE = 4

Products	Distance 4 (Unigram)		Bigram
	AFFIN+CSUIC+Senti140	CSUIC+Senti140	
iPod	53.1	53.6	68.75
CanonG3	36.0	36.0	52.79
Hitachi Router	52.1	52.5	63.77
Norton	31.6	32.8	54.91
Average	49.9	49.4	60

After experimenting with different lexicons with varying window sizes, the results concluded the importance of a good lexicon for sentiment analysis. We further observed that having a rich lexicon like Senti140, the addition of AFFIN was useless as the scores of AFFIN were replicated in Senti140. Therefore, we removed AFFIN from our customized lexical dictionary. When we exploit unigram features, we achieved 49% accuracy with window size = 4

A. Negation

Identifying and handling negation is the most important step in sentiment analysis task. The most challenging part of handling negation is to identify its scope and its effect on sentiment words. In addition, it is not yet cleared that how the effect and resolution of negation should be represented and generalized for different domains [26]. Different experiments showed that the identification and handling of negation improves both accuracy and performance of sentiment analysis system [27]. In our proposed approach, we handled Unigrams and Bigrams with different windows sizes using different lexicons. We have achieved the highest accuracy of 60% with lexicon CSUIC and Senti140 with Window size of four. To handle negation, we used customized list, which contains all the negative words. Previously, for all negative words we were using the score of Senti140 but after writing a separate routine for negative

words, our accuracy shoots up to 10% from previous results. If any opinion word matched with the negative list word, we appended the word NOT to all the surrounding opinion words in the array with the window size of four. For example, “I do not like iPhone5” was converted into “I do not NOT_like NOT_IPhone5.” The average accuracy of algorithm reached to 73.5% after this experiment. To show the effect of negation on sentiment scores, we are multiplying the sentiment scores of words appended with NOT by -1.

TABLE IV
RESULTS AFTER HANDLING NEGATION

Product Name	Accuracy
Norton	73%
Canon G3	85%
Cannon S100	76.25%
Nokia 6600	75.70%
Hitachi Router	74.70%
iPod	76.56%
Average	76.86%

B. Valence Shifter

It is a very challenging task to reflect and distribute the effect of some negative words to its surrounding opinion words. Most sentiment analysis system perform well on majority of text classification problem, but a particular linguistic feature i.e. Valence shifter always poses challenges and problems to these systems. The study showed that almost 15% sentences in reviews contain valence shifters and handling them correctly significantly increases the classification accuracy [28]. Simple example of a review containing valence shifter is “This is not a good book,” but not to our surprise, not many reviews are as straightforward as shown. Especially when consumer put a bad review about any product, they do not express their opinion very directly. For example: “The overly detailed approach makes it a hard book to recommend enthusiastically”, although the reviewer is discouraging the readers to read the book but since the sentence contains words like enthusiastically and recommend, the overall sentiment score of the sentence might get positive score. In our proposed approach, after achieving 77% accuracy, we further enhance our analysis with Negation and run different experiments by changing the multiplication factor.

Previously, some model verbs if found in the opinion array receive their score from either Senti140 or CSUIC lexicon, and some of them were part of the Stop word list. We interchange their score, i.e., if previously some model verbs were getting their score from lexicon; we include those verbs in negative words list so that they could be treated as a negative word; hence, handled by the negation handler routine. If any model verb was part of the stop word list, we remove it and get its score from the lexicon. No model verb shows significant improvement over accuracy when handled individually by adding to the Negative word list or by removing from the Stop Word list. The reason for former is the number of words found in Senti140 lexicon. The dictionary is very large and almost contains the score of every word. Therefore, when we add these model verbs in negative word list, the accuracy does not increase. In contrast, some model verbs like should contribute best if ignore and included in Stop Word List.

C. Diminisher

Diminisher like valence shifter affects the score of preceding opinion words by some factor. Like negative words, diminishers also have considerable effect on sentiment analysis. In our experiment, we have tested some diminisher words and instead of using a multiplying factor of -1 like we have used for negative word list, we use Senti140 score to analyze the results. Following are the results of Diminisher words. The diminishers we used to analyze the changes were *below*, *few*, *over*, *small*, *down* etc.

TABLE V
 RESULTS WHEN SCORE OF SENTI140 WAS USED FOR DIMINISHER WORDS

Product Name	Stop Word List	Using Senti140 dictionary
Norton	72%	73%
Canon G3	84%	83%
Cannon S100	76%	74%
Nokia 6600	77%	76%
Hitachi Router	73%	74%
iPod	77%	76%

V. ASPECT RANKING

After receiving the final sentiment score of top 20 aspects, we rank these aspects' sentiment score using bar graph. The X-axis defined the features or aspect, whereas the Y-axis defined the score. Each product has two separate bar graphs, one for positive, and one negative score aspects.

Figure 2 and 3 shows the experimental results of **Cheetos**. In figure 2, our system has identified positive aspects of *value*, *chips*, and *package* and give score of 1.2, 0.9, and 0.3. Similarly, in figure 3 the negative aspect list contains an important aspect like *price* with score of -0.3. Since the product does not contain much objective reviews on Amazon.com, the system fail to identify some good aspects. Moving further, we run our analysis on two specific smart phone i.e. **Samsung Galaxy S5 and iPhone 6**. The system identified that *fingerprint scanner*, *picture quality*, *battery life*, and *front camera* are the good aspects of S5 (shown in figure 4), whereas *apps* and *pixel density in pictures* are some bad aspects of it. In contrast, the iPhone 6 has *aluminum body*, *design*, and *elegant case* as positive aspect (shown in figure 6) with *button stabilizer* and *mute switch* as negative aspect (not shown in figure). Interpreting the results, customers who prefer better design and body of smart phone to its camera and battery can easily go for iPhone 6, whereas those who prefer good camera with longer battery life can buy Galaxy S5.

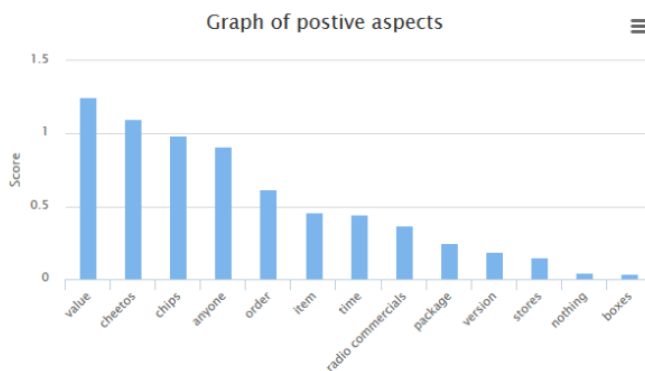


Fig. 2. Positive aspects of Cheetos shown in the form of bar graph

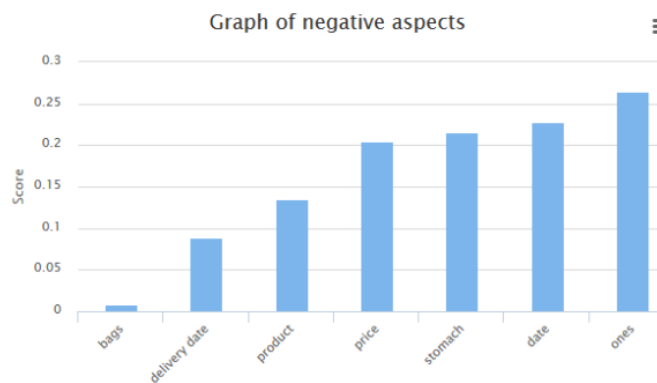


Fig. 3. Negative aspects of Cheetos shown in the form of bar graph

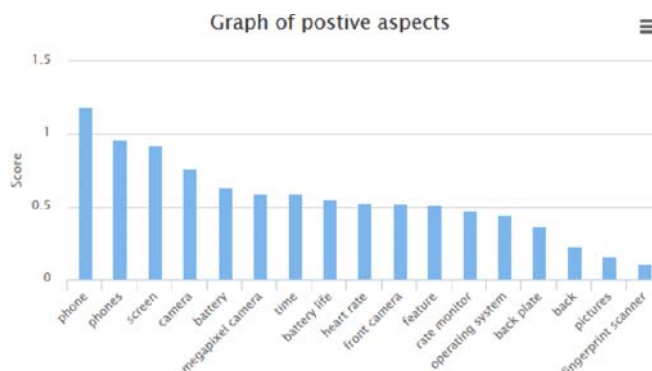


Fig. 4. Positive aspects of Samsung Galaxy S5 shown in the form of bar graph

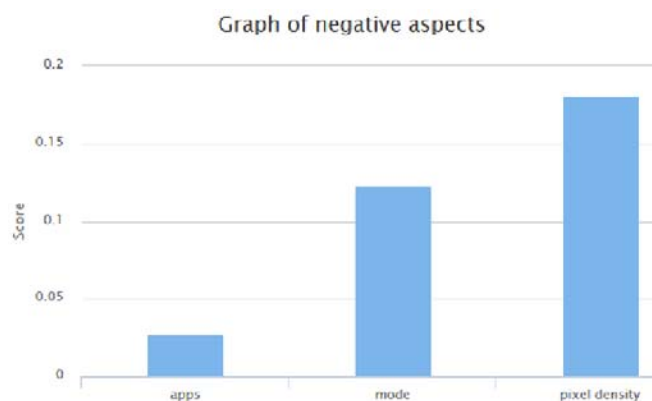


Fig. 5. Negative aspects of Samsung Galaxy S5 shown in the form of bar graph

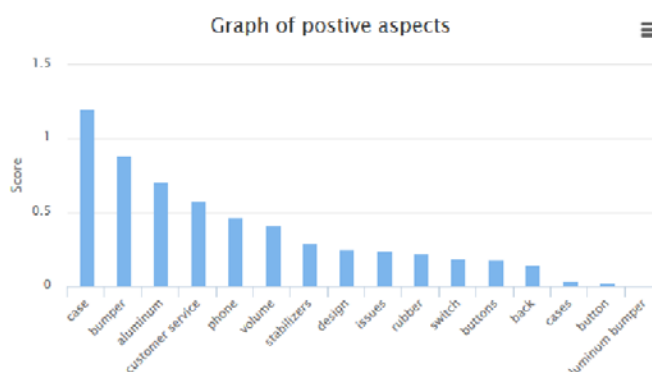


Fig. 6. Positive aspects of iPhone 6 shown in the form of bar graph

The results are dependent on the availability of not only reviews, but also some good reviews. These results are extracted from the reviews of Amazon only, and we are confident that if we increase the dimension of our search, the results will improve. In addition, the results of all the products either technical or non-technical category are

sensitive to what reviews people are giving on Amazon. It is possible that many products do not have good or useful reviews on Amazon, and as a result, the result set may show meaningless aspects.

VI. CONCLUSION

In this research, we have studied and analyzed different factors that could affect the sentiment scores of opinion word. We have studied the importance of having relevant data for the reliability of any experiment and discussed the challenges of extracting target text from un-organized and un-structured spider of web pages. Further, we have discussed the importance of identifying aspects in the problem of product aspect ranking through sentiment analysis. We have proposed a lexicon based approach for product aspect ranking through sentiment analysis. We have exploited the impact of good lexicon on classification accuracy by experimenting with different short and rich lexicons. We also study the importance of handling negation in sentiment analysis task. For handling negation, we proposed a simple and effective approach of making a customized list of negative words and showed the increase in accuracy. Experimental results also proved that if a proper rich lexicon like CSUIC is used for product aspect ranking, then we do not need to handle the valence shifters (VS) and diminishers explicitly. The experimental corpus of review contained 5000 reviews of five different products. We have achieved 78% accuracy in finding the sentiment score of product aspects when used lexicons of CSUIC and Sentiment140 with windows size of 4. We also achieved 77% accuracy on restaurant training dataset of SemEval 2014 with 75% accuracy on Laptop Training dataset. Likewise, our system also achieved 75% accuracy on both trailing dataset of laptop and restaurant. We maintained a customized list of stopping words and negative words, which helped us to deal with negative words more accurately and without isolation.

In future, we will expand our domain from reviews and run our proposed approach in Social media text like Tweets and Facebook status. In addition, we will also experiment with other crude heuristics for identifying aspects from the text instead of targeting nouns only. Likewise, in sentiment analysis, we will expand our opinion dictionary with different other lexicons and analyze its final effect on our current approach. We will also work on neutral sentiment score of aspect and accumulate how to represent an aspect if it has a sentiment score of zero.

REFERENCES

- [1] Anonymous, "Larche Digital Media," Larche, 2014. [Online]. Available: <http://larchedigitalmedia.com/45-awesome-review-digital-marketing-stats-you-should-know/>. [Accessed 12 May 2015].
- [2] M. H. Bing Liu, "Mining and summarizing customer reviews," in *Knowledge discovery and data mining, ACM*, Newyork, 2004.
- [3] J. Y. M. W. T.-S. C. Zheng-Jun Zha, "Product Aspect Ranking and Its Application," *IEEE*, vol. 26, no. 5, pp. 1211-1224, 2014.
- [4] B. P. a. L. L. a. S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques," in *Empirical methods in natural language processing*, Stroudsburg, 2002.
- [5] S. J. Y. D. H. Hanhoon Kang, "Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis," *Expert System with Applications*, vol. 39, no. 5, pp. 6000-6010, 2012.
- [6] S. L. D. M. P. Kushal Dave, "Mining the peanut gallery: opinion extraction and semantic classification of product reviews," in *12th international conference on World Wide Web*, Pasadena, 2003.
- [7] A. B. B. R. D. K. A. M. Michael Wiegand, "A survey on the role of negation in sentiment analysis," in *Negation and Speculation in Natural Language Processing*, Uppsala, 2010.
- [8] M. A. G. Diana Maynard, "Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis," in *Proceedings of LREC*, 2014.
- [9] E. Filatova, "Irony and Sarcasm: Corpus Generation and Analysis Using Crowdsourcing," in *European Language Resources Association (ELRA)*, Istanbul, 2012.
- [10] A. Z. Livia Polanyi, "Contextual Valence Shifters," in *Computing Attitude and Affect in Text: Theory and Applications*, vol. 20, California, Springer Netherlands, 2006, pp. 1-10.
- [11] D. I. Alistair Kennedy, "Sentiment Classification of Movie Reviews Using Contextual Valence Shifters," *Computational Intelligence*, vol. 22, no. 2, pp. 110-125, 2006.
- [12] Z.-J. Zha, "Product Aspect Ranking and its Application," *IEEE "KNOWLEDGE AND DATA ENGINEERING"*, vol. 26, no. 5, pp. 1211-1224, 2014.
- [13] H. a. Liu, "Opinion Mining, Sentiment Analysis, Opinion Extraction," 15 May 2014. [Online]. Available: <http://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html>. [Accessed 9 May 2015].
- [14] D. Ashour, *The Impact of Text Processing and Term Weighting on Arabic Text Classification*, Gaza: The Islamic University, 2010.
- [15] D. Torunoglu, *Analysis of Preprocessing Methods on Classification of Turkish Texts*, Turkey: Dogus University.
- [16] P. Q. Teresa Gonc alves, "Evaluating preprocessing techniques in a Text Classification Problem," in *International Conference*, Portugal, 2005.
- [17] N. Z. a. J. Liu, "The Mining Mart Approach to Knowledge Discovery in Database," *Intelligent Technologies for Information Analysis*, pp. 47-65, 2004.
- [18] R. A. V. Srividhya, "Evaluating Preprocessing Techniques in Text Categorization," *International Journal of Computer Science and Application*, pp. 47-11, 2010.
- [19] S. G. Alper Kursat Uysal, "The impact of preprocessing on text classification," *Information Processing and Management*, vol. 50, no. 1, pp. 104-112, 2014.
- [20] L.-F. C. H.-J. L. Wen-Hsiang Lu, "Anchor Text Mining for Translation of Web Queries: A transitive translation approach," *ACM Transactions on Information Systems*, vol. 22, no. 2, pp. 242-269, 2004.
- [21] Y. C. C. C. J. T. H. Myle Ott, "Finding deceptive opinion spam by any stretch of imagination," in *Human Language Technologies*, Stroudsburg, 2011.
- [22] B. Liu, "Sentiment Analysis Layers," in *Sentiment Analysis and Opinion Mining*, Claypool Publisher, 2012, pp. 10-12.
- [23] S. Muhammad, "Saif Muhammad Homepage," 07 July 2015. [Online]. Available: <http://www.saifmohammad.com/WebPages/ResearchInterests.html>. [Accessed 9 August 2015].
- [24] B. L. Nitin Jindal, "Opinion Spam Analysis," Chicago, Illinois, 2013.
- [25] S. Evaluation, "Metashare," 9 December 2014. [Online]. Available: <http://metashare.ilsf.gr:8080/repository/search/?q=semeval>. [Accessed 9 August 2015].
- [26] E. Lapponi, "Representing and Resolving Negation for Sentiment Analysis," in *Data Mining Workshops (ICDMW)*, Brussels, 2012.
- [27] C. Y. W. M. Lifeng Jia, "The effect of negation on sentiment analysis and retrieval effectiveness," in *18th ACM conference on Information and knowledge management*, New York, 2009.
- [28] L. Dillard, "Sentiment Classification," University of Washington, Washington, 2007.