An Atypical Approach for Uncovering the Essence of Emotions in Consumer Reviews

Sudhir N Dhage[1] and Alfiya Shama[2]
Department of Computer Engineering, Sardar Patel Institute of Technology,
Mumbai, India 400 058.
[1]sudhir_dhage@spit.ac.in, [2]alfiya.shama@spit.ac.in

Abstract—**Reviews stationed on disparate portals of E-commerce and civil networks voice emotions of the consumer experiencing it. This convenient material available works as an absolute sufficient advice demanded by the audience willing to purchase the product or service. Underlying methods discover whether the speaker is satisfied or dissatisfied about a product but fails to imprison the reflection of human emotions. The momentous act of emotions in web reviews and its impact which is often undervalued is explored in this research article. The notion is to detect emotions from customer written text by utilizing the Emotion Revelation Framework (ERF). Analysis of ERF on Customer Review Dataset(CRD) caters to associations on what factors to focus upon for prosperity. An accuracy in terms of F1- score is obtained around 66% making use of CRD dataset.**

Index Terms—**E-commerce; emotions; momentous; Emotion Revelation Framework (ERF); associations; prosperity.**

1. Introduction

E-commerce websites render a stage to consumer community to voice their involvement, acquaintances and proficiency for the products or services they have purchased [15]. The cyberspace boosts consumers to specify their attitudes, sentiment or emotions not restricting to social problems like election results or economical crisis but even for entities they have bought or experienced [2], [13]. Textual content written for products is not something modern but is performed since ages in the form of paper surveys or questionnaire [7]. Before purchasing any product or availing any service people today are hooked to read the reviews written about it in order to make an informed and ethical decision [10].

A high level of diligence towards online content posted is the need of the hour for the marketers and corporation. Since this is a genuine feedback which can be a boon or curse for their business. On a daily basis brand new products or services are being floated in the market [9]. The Internet being available at the finger tips of the world population, folks all over the globe instantly update their part of experience on different conventions irrespective to constraints of time, place, access and its a never ending process [6]. This has left bountiful of information at disposal which has led to information overburden and relevance needs to extracted from this abundance [8].

Sentiments or feelings regarding an entity or affair in the form of textual oriented reviews assists the regular customers to plan their purchasing resolution for future [7], [16]. A lot of researchers have pointed out that sale volumes and impression of brands is troubled by reviews. A reviewer can highlight a lot of aspects about products which may give a nod to buy it [11], [13]. Today exploration is not finite till review content but interest also includes the size, quality of information posted, who has posted it, how regular the person does so, which location is the reviewer residing in and much profound [12].

As soon as an individual makes a purchase of a brand or business, he or she always goes through a spectrum of emotions. When an individual senses any entity he may not be completely happy or completely annoyed but emotions are often conflicted and appear to be mixed [18]. Example of a phone bought, some features like camera quality and memory value may be liked but power backup may be disliked. So emotions do not tend to clung on one side but are often a concoction [14].

Emotions explicitly mention human behaviour towards a concerned element [4], [17]. Emphasis of emotions on sales and profits are appropriately accepted by the marketers and association selling the entity [18]. Studies analogous to sentiment analysis as well as determination of which review is helpful or not helpful has been addressed in the past contrary our aim is directed towards distinctive queries :

- How human emotions can prove to be helpful to associations in their sales and trade ?
- Does particular product features have a greater impact on humans ?
- If there is a change in the entity does it impact the usefulness computed ?
- How many distinct emotions can be attained for a particular product or service ?

The analysis and results obtained augment and correlate to the current studies and research carried out. It contributes to the business associations wherein an approach which helps in determining the act of emotions in reviews written by customers for various features of different products can be constructed [1].

978-1-5386-4744-8/18/$31.00 ©2018 IEEE

The consecutive fractions of the research article is bifurcated as succeeded. Section 2 Related Literature states a quick critique of the articles studied and influenced to fulfill the concerns stated. Section 3 Computative Interpretation provides an analytical perception of the issue giving basic illustrations for improved understanding. Section 4 Emotion Revelation Framework (ERF) administers the entire process in trivial as well as profound steps taken to attain the motivation listed. Section 5 Results and Discussions presents the results obtained in the form of visual and graphical information for a better picture.

2. Related Work

Emotion Recognition is performed in the past by distinct approaches. In article [5], a helpfulness predictive model based on deep neural network was built and the effect of emotions that assist to review helpfulness evaluated. It exposes a method that mines unique and different positive and negative characteristics from text reviews of products by utilizing NRC emotion Lexicon.To add to this, the category of product, user who has reviewed it, clarity, how readable it is, morphology and features that are related to sentiment value are utilized for comparison purpose as well as the helpfulness framework. Experiments have been carried out whose results are based on two real scenario data-sets which indicate that positive features tend to give best results when individual class of aspect is provided. At the same time negative features and clarity features give a precise performance. It is also noticeable that a blend of features tend to provide good influence over helpfulness model. This research gives a good insight as to how emotions can have an impact on consumers, how their emotion values can pull down or promote E-commerce to flourish.

Another proposed work based on emotional behavior is expressed in [3]. The authors wish to build a methodology that can extract the emotional value from the reviews and determine the significance of these emotional aspects within the various classes of products. The method has made use of a lexicon based approach to work on the emotional value of the reviews as well as construct a model to calculate the significance of the emotional aspects based on the review quality. In order to determine whether a review is of good or bad quality it needs to be found out that whether the review has helped someone else in anyway. In this study it has been seen that positive reviews like trust and joy tend to be extremely good for decision making emotional aspects, even though a lot of variation can be detected under classes of products. This approach made use of two kind of dictionaries one was focused on a large vocabulary list another one small one build by an expert. An important shortcoming that could be addressed was to determine what role a reviewer plays in influencing a buyer and prediction of the response of reviewers over the time.

In article [10] a complete framework has been constructed approaching the new form of internet communication, by associating the connected phenomena of client contentment, composition divergence, knowledge query and type of entity to demonstrate the dissemination of written reviews. The study claims to be of significance as the exhaustive literature review on internet communication does not serve the need and can also be conflicting. An investigation of emotional value as well its features in Internet communication have been made. An exploration of emotions from enormous amount of online reviews from the giant E-commerce space Amazon is processed utilizing Natural Language Processing (NLP) methods. Importance of star ratings in reviews have also been evaluated.

Another article working towards emotion analysis but specifically on Chinese reviews is mentioned in [6]. The research is based on user generated reviews based on chinese language on E-commerce websites. Assertions are made that it fills in the issues arising in between different fields of science like Cognition, Computer Science as well as Data Management. It is based upon OCC theoretical model which states a lot of profound knowledge based on human emotion theory as well as cognitive science based human emotions interaction. The extraction of emotions from Chinese reviews by making use of the OCC model which asserts to be utilized for the first time on Chinese product reviews.

3. Computative Interpretation

A lot of miscellaneous factors are involved with the deceived usefulness but this article does not tend to shed light on reviewer's role. It strictly revolves around impact of feelings or attitude. To give a formal meaning to the above stated problem an analytical representation is depicted as follows. A dataset 'D' comprising of summary like reviewer's name, rating given, review posted and other auxiliary details. The reviews comprised in the dataset D are given as :

$$\mathbf{D} = re_1, re_2, ..re_i \tag{1}$$

where 're_i' denotes the reviews available in the dataset. Reviews consists of tokens or words. Tokens grouped together into an array index are adopted as :

$$\mathbf{re} = t_1, t_2, ..t_i \tag{2}$$

where 't_i' denotes different tokens comprised in reviews. The emotions can be bunched together into an array and stated as :

$$\mathbf{EM} = em_1, em_2, ..em_i \tag{3}$$

where 'em_i' denotes different types of emotions that are to be detected and EM is the database where all emotions are contained.

4. Emotion Revelation Framework (ERF)

The approach undertaken to fulfill the laid intentions is executed in a staged manner. It makes use of building rules comprising of word to emotions assignment in manual way rather than making use of the already available lexicons or corpus. This gives a deeper understanding and eliminates out the disadvantages arising from the existing work. The framework as shown in Fig. 1 is mentioned in peculiarity below:

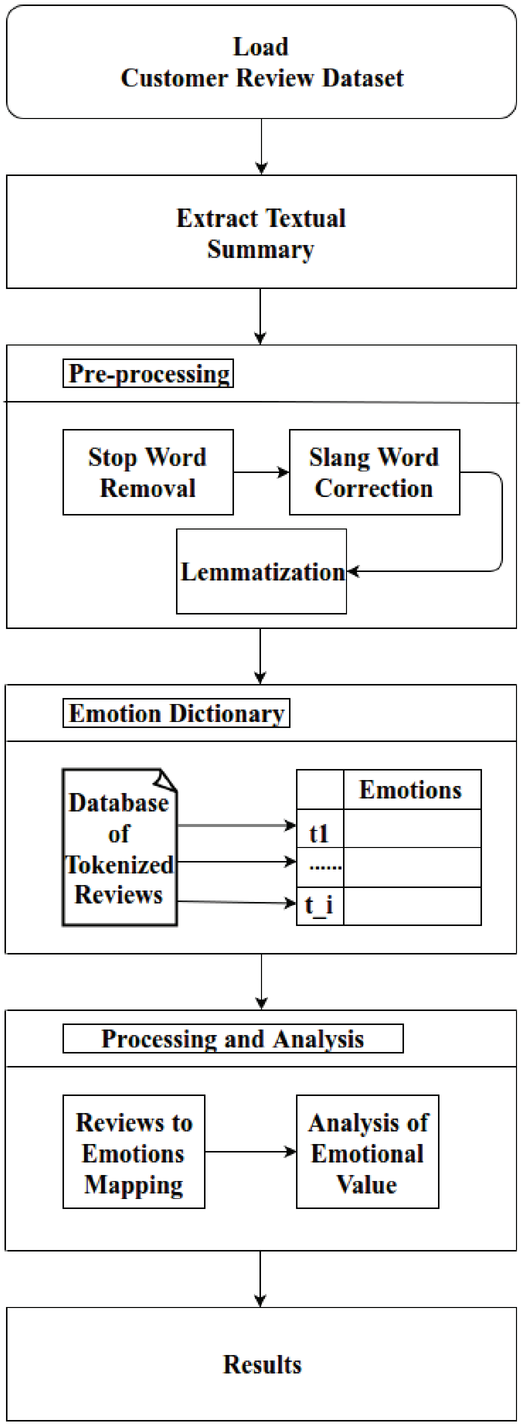
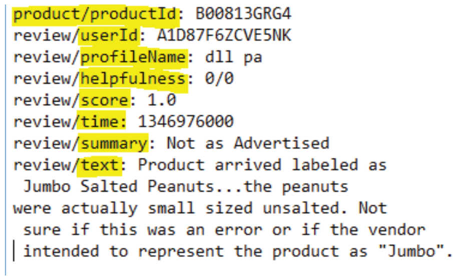


Figure 1. Emotion Revelation Framework (ERF)



```
product/productId: B00813GRG4
review/userId: A1D87F6ZCVE5NK
review/profileName: dll pa
review/helpfulness: 0/0
review/score: 1.0
review/time: 1346976000
review/summary: Not as Advertised
review/text: Product arrived labeled as
 Jumbo Salted Peanuts...the peanuts
were actually small sized unsalted. Not
 sure if this was an error or if the vendor
 intended to represent the product as "Jumbo".
```

Figure 2. Snapshot of a Review comprised in the Customer Review Dataset

4.1. Initial Tasks

Before the prime processing of assignment of emotions a lot of sub tasks need to be performed on the dataset as it is in its raw form for further processing.

Load Customer Review Dataset: Dataset utilized for processing is Customer Review dataset (CRD) which is not restricted to reviews but includes a lot of other information as seen in the snapshot of the dataset posted below (refer Fig. 2). Information posted includes Product id, User id, Name of the reviewer, how helpful the review has been, rating given for the product, at what time the review was set on, the official review written and a short summary highlighting the entire experience. Product ID and Review ID are alphanumeric fields, Profile name is the name set by the reviewer which is a character field, helpfulness and ratings expressed as a numeric field and at the end summary being expressed as a text/character field.

Extract Textual Summary: The main aspiration involves extraction of emotions from reviews posted. Every review comprised in the dataset has a lot of additional information like user id or product id which is of no relevance for the desired aim. A query is run for distilling only the review text in a separate database and filtering out all the other supplementary information.

4.2. Pre-processing

Textual reviews extracted need to be further refined for optimum results. Pre-processing tasks need to applied as mentioned beneath:

Stop Word Removal: Stop words are words that appear to be extremely frequent and natural in any language. They are supportive for building sentences and paragraphs. A lot of ready made lists are available for the English language but still a universally accepted list of words is not yet available. These words do not hold any significance or relevance in identification of emotions therefore needs to be removed. An API that contains a list of stop words is used in the framework which filters out all the stop words available in the reviews by matching it to the bag already available list. A snapshot consisting a few stop words is shown in Fig. 3 for preview.

all	am
an	and
any	are
as	at
be	because

Figure 3. Snapshot of a few stop words

Slang Word Correction: Most of the population today is attracted to social media. In this super paced life people prefer the informal use of words rather than formal. Words like 'ASAP' or 'OMG' are naturally used all over the internet. These words either need to be eliminated or corrected to their grammatical correct forms. An API comprising such words is used for slang words occurrence in reviews for its associated correction to precise English Language words or removal. A slang word list along with it grammatical conversion is presented in Fig. 4

```
cpu#computer
cpy#copy
cr#Can't remember
cr8#crate
crakalakin#happening
crazn#crazy asian
cre8or#creator
crp#crap
crzy#crazy
csi#Crime Scene Investigation
csl#can't stop laughing
```

Figure 4. Snapshot of a few Slang to Grammatical Words

Lemmatization: Lemmatization is used to troop together distinct inflected structures of an altercation, which is called as a lemma. It has its roots from stemming and appears to be similar. Here considerable amount of words are placed under a single roof. A lemmatizer will take different forms of words example bigger, biggest and big all into one umbrella word 'big' as all have the same meaning. Lemmatization is considered to be more cynical as there is a need to understand the background of the word as the procedure is dependent on the word being either a noun, adjective or verb.

4.3. Emotion Dictionary Construction

Reviews are sentences built up using distinct words. To assign emotion to every review separation of words or tokenization needs to be performed on the reviews. Another task is to assign Parts of Speech to very token which denotes which form of token it is at need. The tokens extracted are stored in an array database.

A dictionary is constructed in the form of $[word] \rightarrow [emotion]$ manually. A lot of words from the English Dictionary are seized and by cognitive understanding emotions like joy, sadness, anger, fear, surprise and disgust are assigned to every word. Likewise a database of words and their associated emotions is built and it is called as the emotion dictionary.A snapshot of the Dictionary constructed is available for preview (refer Fig. 5).

```
blue_devils,sadness
bode,fear
boding,fear
bonheur,joy
bore,sadness
bored,sadness
bother,anger
bothersom,anger
bothersome,anger
brokenhearted,sadness
```

Figure 5. Snapshot of the Emotion Dictionary

4.4. Processing and Analysis of Emotions

After having a manually built emotion dictionary at hand the prime part is to assign the various emotions to the numerous reviews posted. A review 're' in dataset 'D' is grabbed. After pre-processing steps a database of the tokens from reviews is extracted. A look up of every token available in reviews is mapped to the emotion dictionary built based on logical rules defined and lastly emotions are assigned to reviews. An example of naive processing of emotions to review mapping is illustrated with a trivial example in Fig. 6. A basic algorithm comprising of rules for emotion assignement are established.

$$N^{joy} = \sum_{w(joy)\epsilon ED} M_w^{w(joy)} count(w(joy))$$

$$N^{sadness} = \sum_{w(sadness)\epsilon ED} M_w^{w(sadness)} count(w(sadness))$$

$$N^{anger} = \sum_{w(anger)\epsilon ED} M_w^{w(anger)} count(w(anger))$$

$$N^{happy} = \sum_{w(happy)\epsilon ED} M_w^{w(happy)} count(w(happy))$$

$$N^{disgust} = \sum_{w(disgust)\epsilon ED} M_w^{w(disgust)} count(w(disgust))$$

$$N^{surprise} = \sum_{w(surprise)\epsilon ED} M_w^{w(surprise)} count(w(surprise))$$

Here, $N^{emotion}$ is the total frequency of the associated emotion occurring in the reviews, w is used to denote the words or tokens in the reviews, ED stands for Emotion Dictionary, $M_{word}^{word(emotion)}$ denotes the weighted occurrence of

the emotion word from the total bag of words and a counter count(word(emotion)) is used to keep a check on the most occurring emotion.
Final equation used to compute the most dominant emotion occurring in the review is illustrated as given below:

$$Emotion = \begin{cases} joy, & \text{if } \frac{N^{joy}}{N^{total}} > Threshold(th) \\ happy, & \text{if } \frac{N^{happy}}{N^{total}} > Threshold(th) \\ sadness, & \text{if } \frac{N^{sadness}}{N^{total}} > Threshold(th) \\ anger, & \text{if } \frac{N^{anger}}{N^{total}} > Threshold(th) \\ surprise, & \text{if } \frac{N^{surprise}}{N^{total}} > Threshold(th) \\ disgust, & \text{if } \frac{N^{disgust}}{N^{total}} > Threshold(th) \\ None, & \text{otherwise} \end{cases} \tag{4}$$

Emotion is predicted based on a threshold (th) set if the value of the total frequency of an emotion divided by the total frequency of all the emotions is greater than a threshold (th) then that emotion is assigned to the review. In case of conflicts the most dominant or the one with maximum value is considered.

5. Results and Discussions

In the final stage, a graphical analysis is performed using the CRD dataset which depicts how emotions are a critical aspect in human cognitivity, usefulness of textual reviews and how humans deeply connect to this form of Internet communication. Analytical methods executed produce diverse and profound statistical results.
7 portrays which emotions were predicted by the framework and how intense they were perceived. Six different emotions namely joy, sadness, fear, anger, disgust and surprise were forecasted by the framework. Amongst these emotion joy appeared to be the most deceived from the Customer Review Dataset (CRD) on contrary fear was seen a least out of all.

Fig 8. depicts after mining every review, most of them hold multiple emotions which conflict with another. Very few appear to be straightforward and produce a single concluding sentiment. For this purpose dominant emotions which means one with maximum value is associated with the review.

Every product or service has certain features which are most spoken about for example a phone, its most common aspects discussed are its weight, camera quality, memory etc. In Fig 9. a distribution of emotions over the distinct categories are visualized. This gives a honest impetus to the associations of the true feelings experienced by customers especially highlighting the peculiar features of the product and not restricted to the product itself.

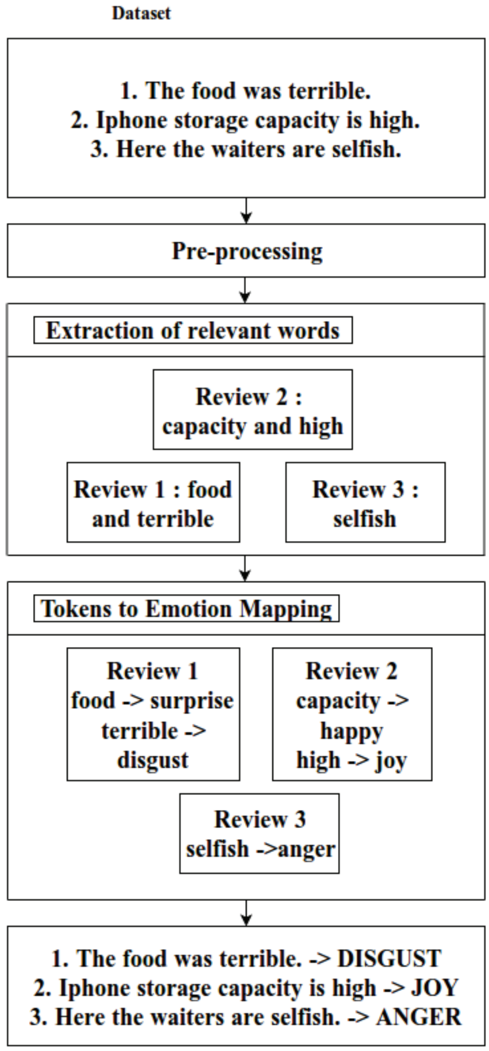


Figure 6. Example Flowchart of Processing of Emotions

6. Conclusion

In recent times there is a change in scenario where marketing associations need not approach people for feedback. There is an epidemic of information overload everywhere. Previous efforts have focused more on polarity identification reviews and exploring which reviews appear to be useful or not fruitful. This research article presents an emotion analysis framework that performs extraction of emotions from reviews posted for various commodities by building an emotion dictionary of feature words. Graphical Evaluations depict different types of significant emotions like joy, sad, fear etc and their relative frequency of occurrence in reviews. Emotions emphasis on various aspects of products by considering an example of an electronic product is visually presented. The framework concludes to analyze human emotions significance in the field of E-commerce, improve consumer's decision making capability by producing intricate and comparative results highlighting

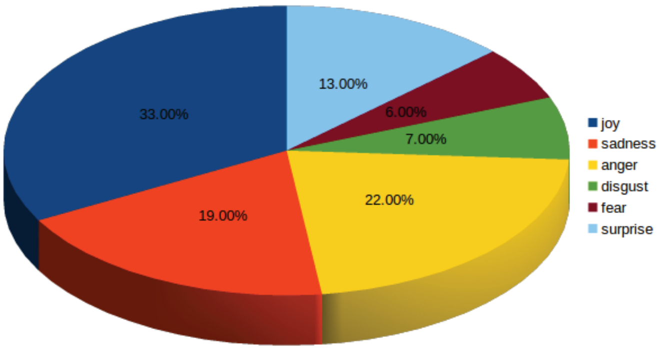


Figure 7. Relative Intensity of Emotions

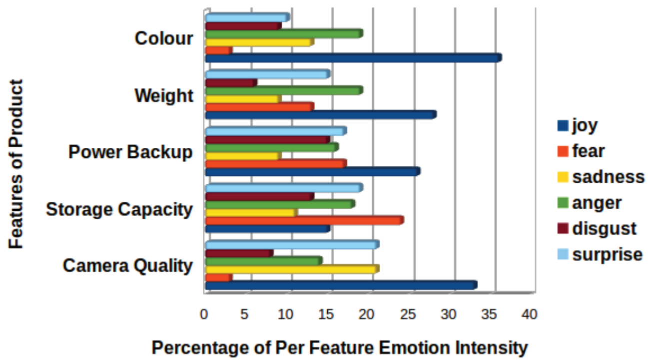


Figure 9. Distribution of Emotions over Product Feature Categories

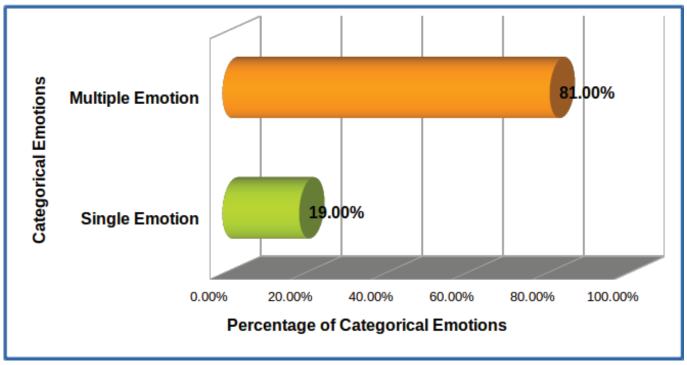


Figure 8. Relative Frequency of Occurrence of Emotions

impact of emotions, boost vendor's revenue as well revision in his strategies from the statistics obtained by motivating them to construct a model using the amount of big data they obtain.

References

[1] J. Kim, P. Gupta, "Emotional expressions in online user reviews: How they influence consumers' product evaluations," *Journal of Business Research,* vol. 65, pp. 985-992, 2012.

[2] V. Priya and K. Umamaheswari, "Ensemble based Parallel k means using Map Reduce for Aspect Based Summarization," in *Proceedings of the International Conference on Informatics and Analytics, ICIA 2016, Pondicherry, India,* ACM, New York, NY, USA, pp. 1-9, 2016.

[3] A. Felbermayr and A. Nanopoulos, "The Role of Emotions for the Perceived Usefulness in Online Customer Reviews", *Journal of Interactive Marketing*, vol. 36, pp. 60-76, 2016.

[4] Y.-H. Hu and K. Chen, "Predicting hotel review helpfulness: The impact of review visibility, and interaction between hotel stars and review ratings," *International Journal of Information Management,* vol. 36, no. 6, pp. 929-944, 2016.

[5] M. Malik and A. Hussain, "Helpfulness of product reviews as a function of discrete positive and negative emotions," *Computers in Human Behavior,* vol. 73, pp. 290-302, 2017.

[6] S. Jiang and J. Qi, "Cognitive Detection of Multiple Discrete Emotions from Chinese Online Reviews," *2016 IEEE First International Conference on Data Science in Cyberspace (DSC),* Changsha, pp. 137-142, 2016.

[7] S. Soni and A. Sharaff, "Sentiment Analysis of Customer Reviews based on Hidden Markov Model," in *Proceedings of the 2015 International Conference on Advanced Research in Computer Science Engineering and Technology ICARCSET '15, Unnao, India,* ACM, New York, NY, USA, pp. 1-5, 2015.

[8] V. Suresh, S. Roohi, and M. Eirinaki, "Aspect-based opinion mining and recommendation system for restaurant reviews," in *Proceedings of the 8th ACM Conference on Recommender systems - RecSys '14, Foster City, Silicon Valley, California, USA,* ACM, New York, NY, USA, pp. 361-362, 2014.

[9] K. Schouten, O. V. D. Weijde, F. Frasincar, and R. Dekker, "Supervised and Unsupervised Aspect Category Detection for Sentiment Analysis with Co-cccurrence Data," *IEEE Transactions on Cybernetics,* pp. 1-13, 2017.

[10] R. Ullah, N. Amblee, W. Kim and H. Lee, "From valence to emotions: Exploring the distribution of emotions in online product reviews," *Decision Support Systems,* vol. 81, pp. 41-53, 2016.

[11] A. Garcia Pablos, M. Cuadros, and G. Rigau, "W2VLDA: Almost unsupervised system for Aspect Based Sentiment Analysis," *Expert Systems with Applications,* vol. 91, pp. 127-137, 2018.

[12] M. S. Akhtar, D. Gupta, A. Ekbal, and P. Bhattacharyya, "Feature selection and ensemble construction: A two-step method for aspect based sentiment analysis," *Knowledge-Based Systems,* vol. 125, pp. 116-135, 2017.

[13] D. Anand and D. Naorem, "Semi-supervised Aspect Based Sentiment Analysis for Movies Using Review Filtering," *Procedia Computer Science,* vol. 84, pp. 86-93, 2016.

[14] B. Gao, N. Hu, and I. Bose, "Follow the herd or be myself? An analysis of consistency in behavior of reviewers and helpfulness of their reviews," *Decision Support Systems,* vol. 95, pp. 1-11, 2017.

[15] M. Fernandez Gavilanes, T. Alvarez Lopez, J. Juncal Martinez, E. Costa Montenegro, and F. J. Gonzalez Castano, "Unsupervised method for sentiment analysis in online texts," *Expert Systems with Applications,* vol. 58, pp. 57-75, 2016.

[16] W. Medhat, A. Hassan and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Engineering Journal*, vol. 5, no. 4, pp. 1093-1113, 2014.

[17] A. Y. Chua and S. Banerjee, "Helpfulness of user-generated reviews as a function of review sentiment, product type and information quality," *Computers in Human Behavior,* vol. 54, pp. 547-554, 2016.

[18] Z. Liu and S. Park, "What makes a useful online review? Implication for travel product websites," *Journal of Tourism Management,* vol. 47, pp. 140-151, 2015.