Analysis of User Review in E-Commerce Websites

Viveka T, M.E[1], Pavithra S[2], Rajeshwari V[2], Sahaya Saranya S[2.]

Teaching Fellow, Department of Computer Science and Engineering, University College of Engineering

Nagercoil,Tamilnadu, India[1]

Department of Computer Science and Engineering, University College of Engineering Nagercoil,Tamilnadu, India[2]

ABSTRACT: Online business is the easiest way of shopping. In online business, users can buy the products by viewing the feedbacks or reviews of the other users who are used the products earlier. Main issue is that after launching the product changes cannot be modified by seeing the reviews. The proposed system is to study the behavior characteristic of early reviewers through their posted reviews on large e-commerce platform. A user who has posted a review in the early stage is considered as an early reviewer. For such online marketing, huge numbers of reviews are given by the users for the products before launching. It is important for companies to identify early reviewers since their reviews can help companies to adjust marketing strategies and improve product designs, which can eventually lead to the success of their new products. As the consumer can able to know the quality of the product, whereas the firm can able to get the feedback of the product. Based on those reviews the other users can able to know what is the good product.

KEYWORDS: Reviewers, NLP algorithm, Reviews

I. INTRODUCTION

Consumer reviews are in valuable as a source of   data to help people form opinions on a wide range of products. Because majority of customers will read online reviews before making an informed purchase. Beyond telling us whether a product is 'good' or 'bad', reviews tell us about a wide range of personal experiences; these include objective descriptions of the products properties, subjective qualitative assessments, as well as unique use cases. We aim to conduct effective analysis and make accurate prediction on reviewers. where people can use natural language rather than keywords to ask questions and seek advice or opinions from real people who have relevant knowledge or experiences. To analyze the characteristics of early reviewers, we take two important metrics associated with their reviews, their [1] review ratings and helpfulness scores assigned by others. We have found that an early reviewer [2] tends to assign a higher average rating score to products; and an reviewer tends to post more helpful reviews.

   Natural Language Processing is a field that covers computer understanding and manipulation of Human Language and its ripe possibility for newsgathering.NLP   is a way for computer to analyze understand, and derive meaning from human language in a smart and useful way by utilizing NLP, developers can organize and structure knowledge to perform task such as automatic summarization, translation, named entity, recognition, relationship extraction sentiment analysis, speech recognition.

        The emergence of e-commerce websites has enabled users to publish or share purchase experiences by posting product reviews, which usually contain useful opinions, comments and feedback towards a product. As such, a majority of customers will read online reviews before making an informed purchase decision [1]. It has been reported about 71% of global online shoppers read online reviews before purchasing a product [2]. Product reviews, especially the early reviews (i.e., the reviews posted in the early stage of a product), have a high impact on subsequent product sales [3]. We call the users who posted the early reviews early reviewers.

II. RELATED WORK

As per our survey, there exist many product that analyse user review using e-commerce website. In website each page has its unique features currently certain companies are officially registered and are working to better and improve there product design. Following description the work being performed by others.

H. D. R, proposed by [1] "MM algorithms for generalized Bradley-terry models," Annals of Statistics, to describe the probabilities of the possible outcomes when individuals are judged against one another in pairs using the Bradley-terry model. T. R. M. J. E. Menke, proposed by [2] "A Bradley-terry artificial neural network model for individual ratings in group competitions," To describe the Neural computing and Applications, to re-parameterizes the Bradley–Terry model as a single layer artificial neural network (ANN) using the Bradley-Terry model. E. Gilbert, [3], proposed by "Understanding deja reviewers." To describe you begin to format your paper, first write and save the content as a separate text file. Complete all content aTo review products on the web invest considerable time and energy. using Margin based embedding ranking model (MERM). R. Peres, [4], proposed by "Innovation diffusion and new product growth models: A critical review and research directions,".To exposed a wide range of influences such as word mouth communication . using diffusion. J. Liu [5], proposed by "Competition-based user expertise score estimation" To estimating the relative expertise score of users in community question and answering services (CQA) using Tapu Search and Support Vector Machine. T. Liu, [6] proposed by," Learning to Rank for Information Retrieval". To Learn to rank is useful for document retrieval, collaborative filtering, and many other applications. using learning method. I. Mele [7],proposed by "The early-adopter graph and its application to webpage recommendation," To identify users who discover interesting pages before others. Using early-adopter graph. K. Sarkar [8], proposed by "How do we find early adopters who will guide a resource constrained network towards a desired distribution of behaviors". To distribute the behaviours amongst the early adopters to achieve a target distribution in the population. using Non deterministic Polynomial time(NP complete). J. McAuley [9], proposed by "Addressing complex and subjective product-related queries with customer reviews," To automatically learn whether a review of a product is relevant to a given query. using Relevance ranking.

III.ARCHITECTURE DESCRIPTION

The proposed system has three main parts.Company Admin and User.First the company can post the product and description about the product.Second the user can view the product details and give the reviews.and Finally the admin can analyze the users review that analyzing details are given to the company.
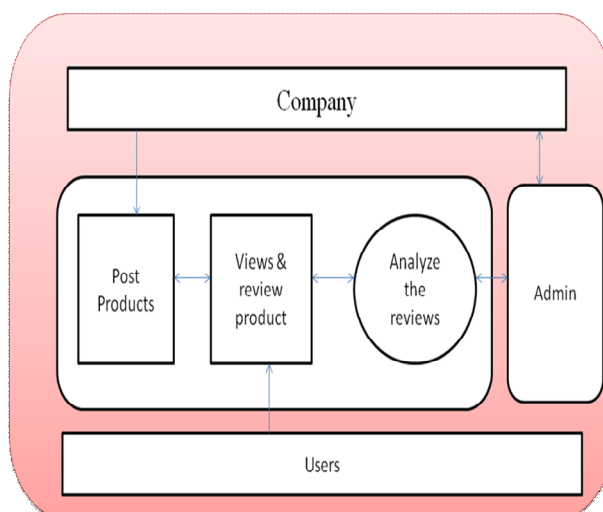

Fig:3.1.Architechture diagram

IV.PROPOSED METHADOLOGY

The main objective of this paper is to analyze the user reviews before launching the product. The companies are register in our websites. and they are posting their product features. Then the user login our website and view the product and posting the reviews. Then the admin analyze the user reviews using Natural Languages Processing(NLP) Algorithm. And finally the admin give above analyzed review to the company. since their reviews can help companies to adjust marketing strategies and improve product designs, which can eventually lead to the success of their new products. As the consumer can able to know the quality of the product. The following step are how to work the NLP algorithm.

STEPS OF NATURAL LANGUAGE PROCESSING
There are 5 phases involved in natural language processing
1. Morphological and Lexical Analysis
2. Syntactic Analysis
3. Semantic Analysis
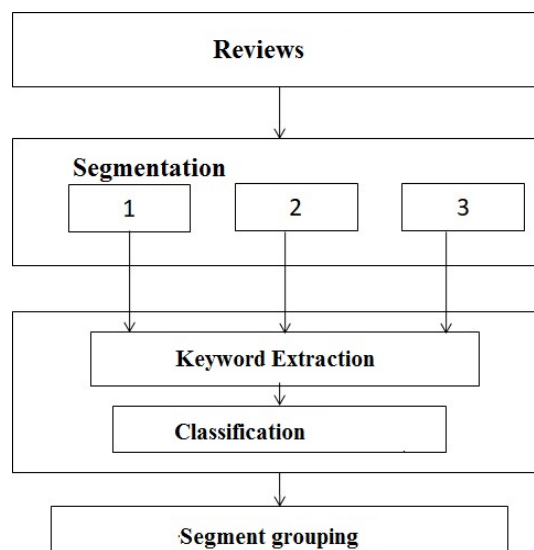4. Discourse Integration
5. Pragmatic Analysis

Fig:4.1 Detailed design

1 Morphological and Lexical Analysis
The lexicon of a language is its vocabulary that includes its words and expressions. Morphology depicts analyzing, identifying and description of structure of words.
1.1 Lexical Analysis
It involves dividing a text into paragraphs, words and the sentences

2 Syntactic Analysis
This involves analyzation of the words in a sentence to depict the grammatical structure of the sentence. The words are transformed into structure that shows how the words are related to each other eg. "the girl the go to the school". This would definitely be rejected by the English syntactic analyzer.

3 Semantic Analysis

This abstracts the dictionary meaning or the exact meaning from context. The structures which are created by the syntactic analyzer are assigned meaning. There is a mapping between the syntactic structures and the objects in task domain.eg "colorless blue idea" .This would be rejected by the analyzer as colorless blue do not make any sense together.

4 Discourse Integration

The meaning of any single sentence depends upon the sentences that proceeds it and also invokes the meaning of the sentences that follow it.eg the word "it" in the sentence "she wanted it" depends upon the prior discourse context.

5 Pragmatic Analysis

It means abstracting or deriving the purposeful use of the language in situations importantly those aspects of language which require world knowledge the main focus is on what was said is reinterpreted on what it actually means. eg "close the window?" should have been interpreted as a request rather than an order
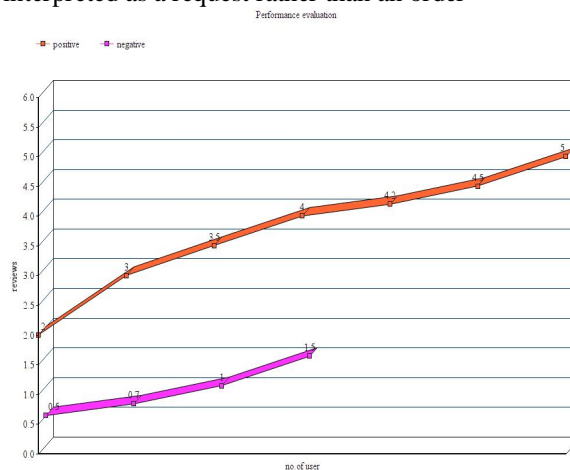

Fig:4.2 Review analyzed graph

NLP algorithm:
Step 1:
capture unique stemmed reviews in the training corpus
        corpus_words = {}
        class_words = {}
Step 2:
tokenize each sentence into words
        for word in nltk.word_tokenize(data['sentence']):
Step 3:
turn a list into a set (of unique items) and then a list again (this removes duplicates)
        classes = list(set([a['class'] for a in training_data]))
        for c in classes:
Prepare a list of words within each class
        class_words[c] = []
Step 4:
ignore a some things
        if word not in ["?", "'s"]:

```
stem and lowercase each word
stemmed_word = stemmer.stem(word.lower())
Have we not seen this word already?
if stemmed_word not in corpus_words:
 corpus_words[stemmed_word] = 1
else:
        corpus_words[stemmed_word] += 1
```

Step 5:
```
 # we have all words in each class
 print ("Class words: %s" % class_words)
```

Step 6:
```
return the class with highest score for sentence
def classify(sentence):
        high_class = None
        high_score = 0
   # loop through our classes
for c in class_words.keys():
     # calculate score of sentence for each class
score=calculate_class_score_commonality(sentence,c, show_details=False)
     # keep track of highest score
if score > high_score:
        high_class = c
        high_score = score
return high_class, high_score
```

V.SCOPE AND APPLICATION

An early reviewer tends to assign a higher average rating score. An early reviewer tends to post more helpful reviews. Our analysis of product reviews also indicates that early reviewers' ratings and their received helpfulness scores are likely to influence product popularity. The user can able to know what is the good product. It is important for companies to identify early reviewers since their reviews can help companies to adjust marketing strategies and improve product designs, which can eventually lead to the success of their new products. . As the consumer can able to know the quality of the product.

VI.CONCLUSION

The prime objective of our project is analyze the user reviews. The reviews are gathered from user, when before the product has been launched in the market. Our empirical analysis strengthens a series of theoretical conclusions from sociology and economics. We found that (1) an early reviewer tends to assign a higher average rating score; and (2) an early reviewer tends to post more helpful reviews. Our experiments also indicate that early reviewers' ratings and their received helpfulness scores are likely to influence product popularity.

        .NLP is a way for computer to analyze understand, and derive meaning from human language in a smart and useful way by utilizing NLP, developers can organize and structure knowledge to perform task such as automatic summarization, translation, named entity, recognition, relationship extraction sentiment analysis, speech recognition. Other algorithms can use different methods to doing the above processes. But NLP algorithm doing all the above processes .

    . Based on those reviews the other users can able to know what is the good product. So NLP algorithm can reduce the time and save the cost.

REFERENCES

[1]   S. Chen and T. Joachims, "Modeling intransitivity in   matchup and comparisonindata," in   WSDM, 2016, pp.227-236

[2]   E. Gilbert and K. Karahalios, "Understanding deja   reviewers." in CSCW, 2010, pp. 225–228

[3]   J.Liu, Y.-I. Song, and C.-Y. Lin, "Competition-based user expertise score estimation," in SIGIR, 2011, pp. 425–434

[4]   T. Liu, Learning to Rank for Information Retrieval. Springer, 2011

[5]   J. McAuley and A. Yang, "Addressing complex and subjective product-related queries with customer reviews," in WWW, 2016, pp. 625–635

[6]   I.Mele, F. Bonchi, and A. Gionis, "The early-adopter graph   and its   applicationtowebpagerecommendation," in CIKM, 2012, pp. 1682–168T

[7]   R. M. J. E. Menke, "A bradleyterry artificial neural network model for individual ratings in groucompetitions," Neural computing and Applications, vol. 17(2), p. 175186, 2008.

[8]   R.   Peres,   E.   Muller,   and   V.   Mahajan,   "Innovation   diffusion   and   new productgrowthmodels:Acriticalreviewandresearchdirections,"InternationalJournal of Research in Marketing, vol. 27, no. 2, pp. 91 – 106, 2010.

[9]   K . Sarkar and H. Sundaram, "How do we find early adopters who will guide a resource constrained network towards a desired distribution of behaviors?" in CoRR, 2013, p. 1303.

[10] H. D. R, "Mm algorithms for generalized bradley-terry models," Annals of Statistics, pp. 384–406, 2004