A Proposed Framework for Measuring Customer Satisfaction and Product Recommendation for Ecommerce

Sree Lakshmi Addepalli
Department of Computer Engineering,
VESIT, Mumbai

Sree Gowri Addepalli
Department of Computer Engineering,
VESIT, Mumbai

Monica Kherajani
Department of Computer Engineering,
VESIT, Mumbai

Harsha Jeshnani
Department of Computer Engineering, VESIT, Mumbai

Sujata Khedkar
Associate Professor
Department of Computer Engineering, VESIT, Mumbai

ABSTRACT
Ecommerce is a significant part in the global economy. Ecommerce websites users have high suppositions for the quality of products or services it provides, and if they are not met, the next ecommerce site is only a click away. Managing customer satisfaction, trust and loyalty of ecommerce services is very important for the long-term growth of many businesses. The tremendous growth of consumer-generated content on these sites has led to the development of big data analytics to solve real-life problems. This study aims to examine and demonstrate the utility of big data analytics to understand various important ecommerce-issues, namely the relationship between customer product reviews and satisfaction. Specifically, this study acknowledges semantic analysis and machine learning approach to a large quantity of consumer reviews to deconstruct consumer product reviews and examine its association with satisfaction ratings. This paper studies various approaches to build a framework for an ecommerce which is capable of assessing and modeling satisfaction of customers as well as deducing business intelligence for managers from online product reviews.

Keywords
Ecommerce, Customer satisfaction, Business Intelligence, Big Data Analytics, Product Reviews, Framework, Customers, Managers, Features, Semantic Analysis, Sentiment Score, Machine Learning, Recommendations, Visualization, Metrics.

1. INTRODUCTION
How are products purchased online nowadays? The first thing any customer does is to look for the following things on an ecommerce site:

i. Select the product according to required features. E.g. metallic body for the laptop or camera resolution of a smartphone.

ii. Assess the customer reviews of that product present on site, before taking a decision.

iii. Consider the recommendations given on the ecommerce site.

Cheong and Morrison [1] carried out interviews with participants and studied the impact of online recommendation on consumers' decision. It would have been a good initiative if a small system in the ecommerce platform were built which could actually measure the quantitative and qualitative aspects of the product instead of just showcasing the products. For example consider this particular product review:

"I like Micromax's multimedia features but the battery life sucks".

This sentence has a mixed emotion. Emotion regarding multimedia is positive whereas that regarding battery life is negative. Hence it is important to extract only those opinions relevant to a particular feature and classify them instead of taking the complete sentence and the overall sentiment.

With the advent of Big Data Analytics and Machine learning, it has been possible to make systems intelligent, given them the power to analyze texts and derive insights which could be beneficial for businesses. In Ecommerce platforms with visitors hitting up to 2 billion, there is a lot of raw data that can be analyzed to understand the consumer behavior. If the user clicks, reviews (30 million), tags, blog posts, ratings all are gathered and processed many form of indirect indicators can be converted to key performance indicators of a business thereby helping businesses to find target areas to be profitable from manager perspective. The customers are the sole purpose of any business; measuring their satisfaction level through these feedbacks can help both the customers and managers. Visuals for long have been making a long lasting impact on individuals, thereby empowering them in making quicker decision. Providing them with visualizations and recommendations on the basis of what others are saying can be quite useful.

2. RELATED WORK
Piyush Gupta et al. [2] performed sentiment analysis on product reviews using Hadoop Distributed File System (HDFS) that stored data and used MapReduce architecture for processing. The proposed method worked in two phases: (1) Mapper mapped input key/value pairs to a set of intermediate key/value pairs. It consisted of File Parser (split data into records) and Information Extractor (used to get key/value pairs that were necessary to do sentiment analysis). (2) Reducer performed sentiment analysis of input file.

It worked in three steps: (i) Stop word extraction (ii) Negation Checker: Reversed the polarity of the word whenever it is preceded by a negation. (iii) Polarity Calculation: Identified whether the sentence is positive or negative. Calculated both positive and negative polarities and

after subtracting these polarities determined the overall polarity of the sentence.

Prashast Kumar Singh et al. [3] automated the process of gathering user review data for a particular product from any ecommerce platform and analyzed those reviews in terms of the sentiments expressed about specific features of the product. It involved: filtering of unwanted reviews; tokenizing the reviews into set of sentences, which were further broken down to a set of tags of which later dependencies were calculated like that of an adjective and noun; Quantitative analysis of sentiment by extracting tags that store sentiment and scoring them on the basis of implication; Visualizing the sentiment ratings and the quantitative results generated in the intermediary steps on the basis of features. These results could be used to improve business outcomes and ensure a very high level of customer satisfaction.

Siddharth Jain et al. [4] defined context specific grammar as semantics for a movie domain where features obtained from matching semantic patterns were used to predict the sentiment polarity and give a sentiment score. The proposed methodology assigned a sentiment score to the reviews on the basis of number of matches against the predefined positive and negative grammars. Feature Extraction was done to make the machine

learn the characteristics of a review so that it can classify it accurately as positive or negative. He built a machine-learning model having three stages, which include vectorizing, training and prediction for mapping feature with sentiment score and they used the trained models to predict sentiment polarity of reviews in the testing dataset. The proposed model was flexible and can be applied to any domain by redefining the semantics specific to that domain.

Bing Liu and Minqing Hu [5] summarized all the customer reviews of the product. This paper primarily focused on mining opinions and product features. The two main steps used for summarization of reviews: feature extraction and opinion direction identification. First, they generated frequent features using Apriori algorithm and after pruning out irrelevant features, they identified opinion direction to summarize the opinions of features as positive or negative.

Dr. M P. Singh et al. [6] implemented two parse algorithm for automatic product review analysis using weighted KNN classifier which was successful in segregating weakly and mildly polar reviews as opposed to highly polar reviews. For the first parse algorithm, three files were used: the test dataset, the positive review file and the negative review file.

Table 1: Gap Analysis

SR. NO.	AIM	TECHNIQUE (Precision %)	TEHNOLOGY	DATASET	GAP IDENTIFIED
1.	Analyze the product reviews using Hadoop Distributed File System (HDFS) for storing data and Map Reduce Architecture for performing sentiment analysis [2]	Map Reduce	Hadoop, SentiWordNet for sentiment scoring	Amazon Product Review Dataset	Calculates the overall sentiment polarity of sentence. Does not perform feature based sentiment analysis
2.	Automate gathering of user review data from ecommerce platform and analyze those reviews in terms of sentiment expressed about specific features [3]	NLP based algorithms	MongoDB as NoSQL storage, Stanford NLP parser, SentiWordNet for sentiment scoring	Scraped Review Dataset	It does not present temporal analysis of reviews which can enable manufacturers to look at the sentiments as a function of time
3.	Define context specific grammar as semantics for a movie domain where features obtained from matching semantic patterns are used to predict the sentiment polarity and give a sentiment score [4]	Machine Learning classifiers like Naive Bayes (60%), Linear SVM (63.58%), Sequential Minimal Optimization (62.54%) and Logistic Regression (63.38%)	Apache Spark, HDFS, PySpark Python API, Weka	IMDB Movie Review Dataset	The results of this type of analysis is useful only for business managers and not for customers
4.	Mining opinion and product features for summarizing online reviews [5]	Frequent feature identification using Apriori algorithm, POS tagging	WordNet to determine opinion sentence orientation	Crawled Dataset of Reviews from Amazon	They have not performed clustering of synonymous features
5.	Supervised semantic analysis of product reviews using weighted kNN classifier [13]	Weighted kNN algorithm for classification of reviews	Python 3.2 with NLTK, WordNet	Crawling ecommerce sites for review collection	It uses unigram model (single word, no phrases). Thus, it can only classify sentiment words but cannot associate those sentiments with the specific feature of product

After the first parse the output had a main file of review id, review and corresponding positivity score determined by the algorithm; a file each of positive, negative and weakly polar reviews was returned. After that, the output files of first parse were used as input to the classifier for learning, which was processed on the weakly polar reviews to give scoring to mildly polar reviews. The output of the second parse was a set of reviews with positivity score. The drawback of using this approach was that it used unigram model, which considered words as independent of each other and hence this approach would be suitable only for classification of sentiments but would not be apt for feature based sentiment analysis.

Sristi Khanna et al. [7] proposed a system that was capable of filtering a large corpus of reviews on the basis of its helpfulness rating. The system consisted of three main models: Predictive, Explanatory and Feature model. They adapted a two level filtering approach (Predictive and Explanatory). Predictive model predicts the helpfulness of reviews from the helpfulness ratio by making use of decision trees. In explanatory model, the helpfulness of filtered reviews obtained from predictive model is determined using metrics like subjectivity, readability score, etc. Feature model extracts the features and their related sentiments thus providing a detailed analysis of product performance.

Shu Zhang et al. [8] proposed a recommendation algorithm, which refined the user ratings data to reduce the number of random variables under consideration, and then used a new similarity measure to find target user's neighbor. After the neighborhood of user was found, the weighted average of neighbors' ratings was calculated weighted by their similarity to target user.

Arjit Sachdeva et al. [9] provided an analysis platform, which summarized positive and negative reviews according to product features for a manager. It used Natural Language Processing for tokenization of reviews into set of sentences, which were further broken down to a set of tags of which later dependencies were calculated like that of an adjective and noun. The system believed in visualizing the sentiment ratings for each feature of the product on the basis of its sentiment score.

3. MOTIVATION

The overall rating of the product given by ecommerce users does not provide other customers with a clear understanding of the user's perspective on feature wise performance of the product. Also, business managers do not get an accurate insight of how these ratings affect the sales of the product. Although several methods have been proposed to use product reviews for business intelligence, limited work has been done on using customer satisfaction as a metric. Prior works had customer satisfaction ratings and recommendations independent of each other. Hence the focus is on building a framework for measuring customer satisfaction which can help the business administrators to take strategic decisions as well as provide the customers with valuable infographics which would help them in taking informed decisions regarding which product should be purchased and give them recommendation on the basis of the generated satisfaction rating. The aim is to extract patterns and develop them into insights based on aspect based sentiment analysis on product reviews aggregated from the ecommerce sites and present solutions by visualizations for the management and satisfaction based product recommendation for customers so as to enhance transactions on these sites.

4. EXISTING APPROACHES

The process of mining the opinions or views of the users for the products or services they have used and detecting the orientation of the sentiment of the sentence is sentiment analysis. Sentiment analysis means to infer the opinion polarity of the review i.e. deciding whether the opinion expressed is positive or negative or neutral. Applications of sentimental analysis are found in product reviews, stock market prediction, election results predictions and political debate analysis. Sentiment analysis can be considered as a classification process divided into the different levels of study as: document-level, sentence-level, aspect/feature-level sentiment analysis.

Document-level Semantic Analysis: In this approach of sentiment analysis, the whole document of the total reviews for a particular product or service is considered as the basic input information for deciding the polarity of the sentiment or opinion of the reviewer.

Sentence-level Semantic Analysis: Sentence level classification of sentiments considers the overall opinion of the review text. Often, this type of classification is insufficient because it does not assign sentiments to targeted features of the entity (product or service).

Aspect/feature-level Semantic Analysis: This type of sentiment analysis is used to identify different aspects of a particular entity (product or service) from the reviews and then associate each of these aspects with its corresponding sentiment polarity. It consists of two main tasks:

i. Feature/Aspect Extraction: An opinion word i.e. the quality aspect review is always associated with a target attribute of the entity being reviewed. Target refers to the topic to be extracted from a sentence.

ii. Aspect Sentiment Classification: Aspect sentiment classification refers to the process of deciding whether the aspect of the entity (product or service) can be said to have a positive or negative or neutral sentiment.

Sentiment analysis is a research area, a subset of text mining, which is gathering attention from different domains like information extraction, natural language processing, machine learning, and data mining.
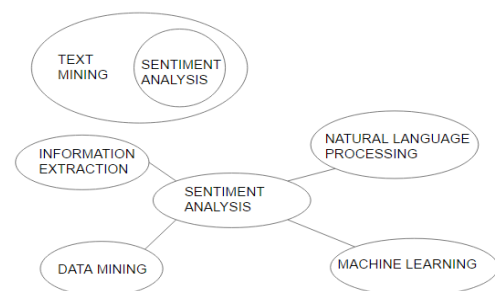

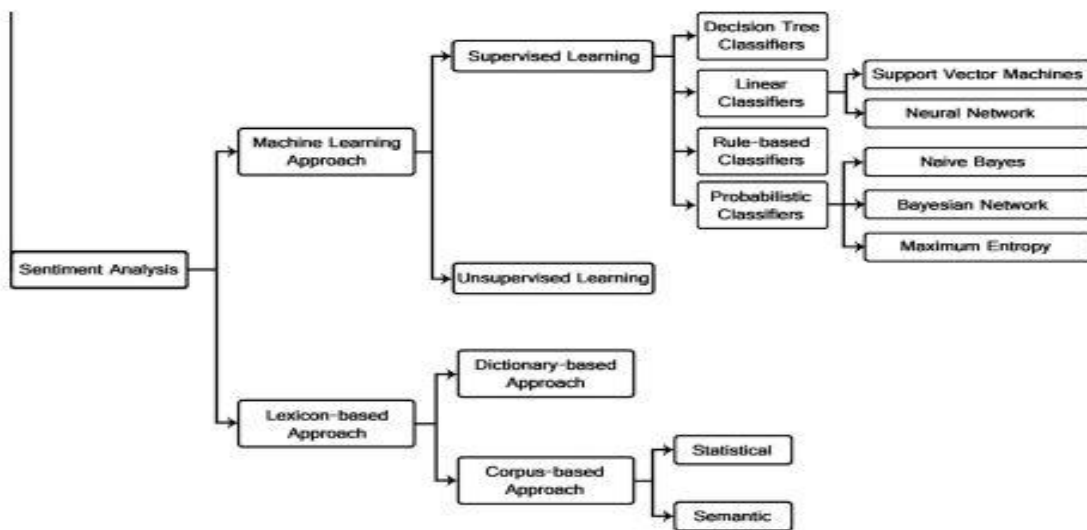**Figure 1: Domain of Sentiment Analysis**

5. CLASSIFICATION TECHNIQUES

Figure 2: Approach to Sentimental Analysis

5.1 Machine Learning

5.1.1 Supervised Learning:

Supervised learning refers to a class of classification algorithms, which make use of labeled training data for building a predictor model. The model could then be used for making reasonable predictions for new input data. Various algorithms that come under this category are summarized in the subsequent sections.

5.1.1.1 K-nearest neighbor classifier (KNN)

KNN is an instance based learning machine-learning algorithm, which chooses the K-nearest neighbor's from the training set and classifies an unannotated input from testing set based on similarity between these K-neighbor's. KNN is used to test the similarity between documents.

5.1.1.2 Naïve Bayes

It is a probabilistic classifier. The main idea of naive Bayes classifier is to calculate the probability that a sentiment word belongs to a particular class by using Bayes rule in probability theory.

P(label|features) = P(label)*P(features label)/p(features)

In sentimental analysis the labels are usually positive, negative or neutral and features extracted are classified in these respective labels. Assuming that all features are independent, the equation could be rewritten as follows:

P(label|features) = P(label)*P(f1|label)*P(f2|label)*

P(fn|label)*/p(features)

5.1.1.3 Support Vector Machines (SVM)

It is a linear classifier where if X ={x1 ......: xn} is the normalized document word frequency, vector A = {a1 ......

an} is a vector of linear coefficients with the same dimensionality as the feature space, and b is a scalar; the output of the linear predictor is defined as p =A.X + b, which is the output of the linear classifier. The predictor p is a separating hyperplane between different classes.

5.1.2 Unsupervised Learning:

Unsupervised learning refers to a class of classification algorithms dealing with unlabeled data. It is based on drawing inferences by observing the recurring patterns in the dataset available & then making use of such inferences for classifying the new input data. It mainly involves clustering and classification techniques.

5.2 Lexicon based Approach:

In this methodology phrases and idioms forming opinions are used for analysis and hence called opinion lexicons. These lexicons are further classified as positive negative or neutral. It usually undergoes text analytics and the sentiment is mostly an adjective and feature is considered a noun. So these adjective noun phrases tend to be opinion lexicons.

5.2.1 Dictionary based approach:

In this approach a previous list of words are used to map with the extracted sentiment. The list keeps growing by searching the synonyms of the words. The disadvantage here is the dictionary may not be context specific or related to the domain hence the opinion word extracted might not be very accurate.

5.2.2 Semantic based approach:

Here the sentiment score is directly predicted by using similarity between words computed by different algorithms. The similarity between words is calculated by its semantic relatedness, which entitles them with the same score. So, for example, if an adjective has a score of two, in the opinion lexicon as mentioned above the noun too gets the score.

6. FEATURE EXTRACTION METHODS

Features are usually nouns about which people give opinions about. The following are some feature extraction techniques:

6.1 Comparison with predefined domain

This technique is used when the domain knowledge is present that is there is a predefined list of features and its synonyms, in this method the feature only needs to be matched with the data.

6.2 Machine Learning Techniques

Machine learning techniques like support vector machine can be applied to train the program to identify the features. But this makes the process too complex as for sentiment

classification machine learning is being used and another dataset would be required for this purpose.

6.3 Noun Retrieval and Processing

This method consists of extraction of nouns from the data to list the potential features. Techniques involved in this step are retrieving list of nouns in data, irrelevant noun removal, frequent nouns identification, relevant nouns identification, feature indicator words mapping and unrelated nouns removal using search engines.
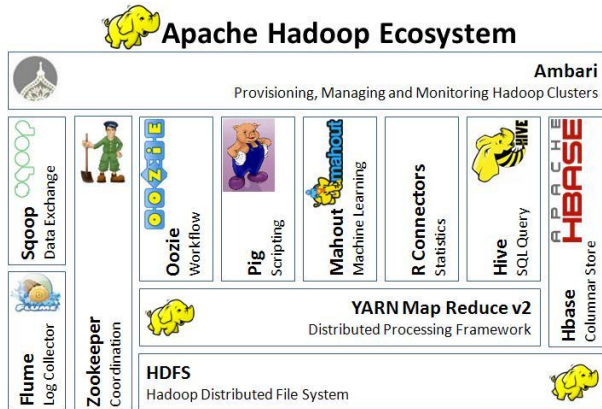
7. SENTIMENT ANALYSIS IN HADOOP

Figure 3: Hadoop and its Ecosystem

This Hadoop is Apache's open source project for analyzing big data both in structured and unstructured format. Hadoop has various components like HDFS - the hadoop file system, SQL query like tool called hive to interact with the data, Apache Hbase, a NoSQL database for unstructured format, Apache mahout a machine learning library and Apache storm for real time data analysis. Hence this tool is perfect for sentimental analysis as it even provides distributed computing environment, which helps from data loss failure and speeds up the processes by parallel computing using Map-Reduce paradigm. It is composed of a Map () procedure that performs filtering and sorting and a Reduce () method that performs a summary operation.

8. RESULTS AND DISCUSSIONS

A survey conducted on a sample of e-shoppers and the literature survey formed a strong base to support this research. Various dimensions were put under scrutiny through the series of questions. The following figures show the results of the survey.
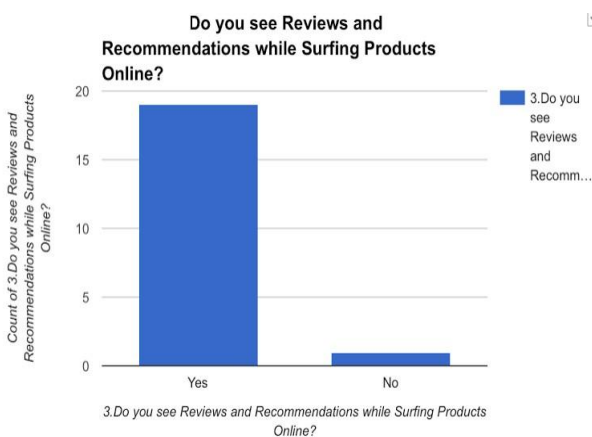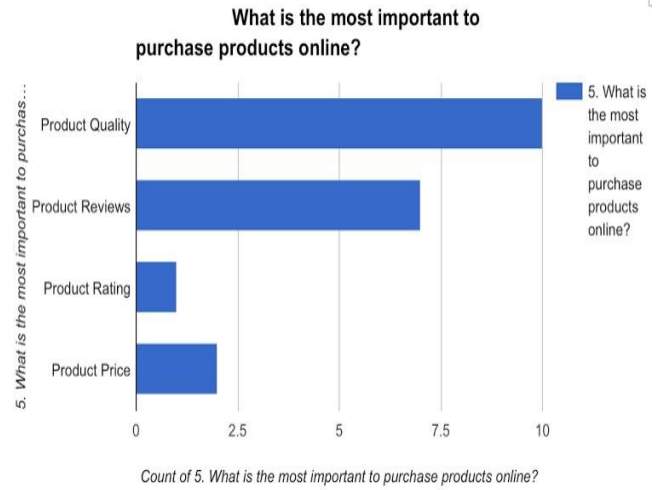

Figure 4: Survey Question 1
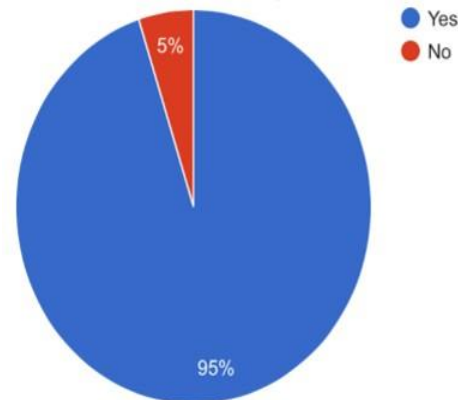

Figure 5: Survey Question 2

Figure 6: Survey Question 3

The conclusions made show that there is a need for the following:

1. Extraction of features from product reviews.

2. Provision of sentiment mapping and sentiment scores to the features with the help of a pre-determined dictionary.

3. Recommendations to the customers based on sentiment of the features.

4. Visualization for managers and customers for enhancement of business strategies at different levels.

5. Provision of dashboards to the managers to get an insight of the sales and performance of the product based on its feature.

9. CONCLUSION AND FUTURE WORK

Concluding with the above research we have decided to build a framework, which extracts a list of features from each review of different products after processing and obtains corresponding sentiment scores. Based on these scores, the customer's satisfaction of each feature of the product can be compared and used for the manager's

benefits by helping him to identify profitable target areas and providing recommendations to customers. Individual measures may be scaled for simple comparison, and combined to give an overall satisfaction rating. It also visualizes the expressed sentiments in way of intuitive and easy to understand graphs, charts to facilitate end user (customers/managers) in taking quick decisions. Hence creating this framework can help predict the satisfaction of all customers for particular features of a product, thus providing a measure of the market's perceived quality of one product model relative to another. The following diagram is the proposed architecture:
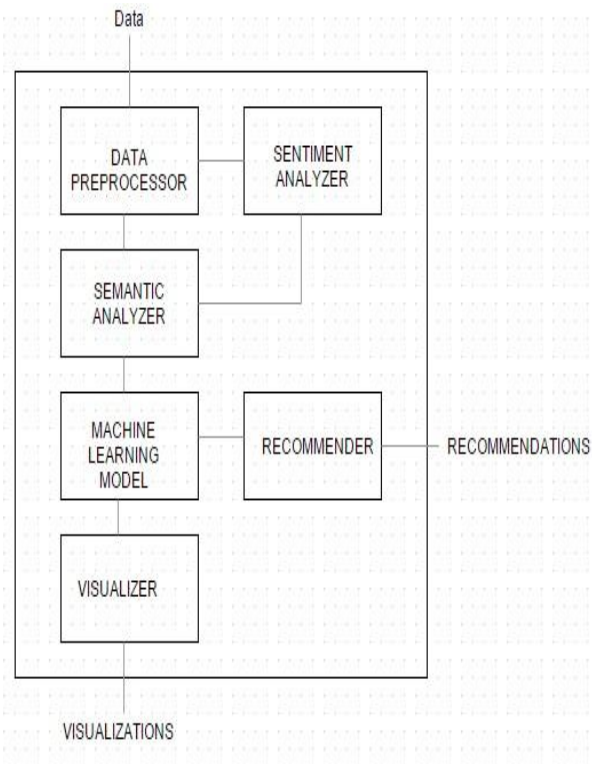

Figure 7: Proposed Framework Design Diagram

10. ACKNOWLEDGEMENT

11. REFERENCES

[1] Cheong, H. J., & Morrison, M. A., "Consumers' reliance on product information and recommendations found in UGC" in Journal of Interactive Advertising, 8(2), pp. 38-49, 2008.

[2] Piyush Gupta, Pardeep Kumar, Girdhar Gopal, "Sentiment Analysis on Hadoop with Hadoop Streaming", International Journal of Computer Applications (0975 – 8887) Volume 121 – No.11, July 2015.

[3] Prashast Kumar Singh, Arjit Sachdeva, Dhruv Mahajan, Nishtha Pande, Amit Sharma "An approach towards feature specific opinion mining and sentimental analysis across e-commerce websites", 5th International Conference- Confluence The Next Generation Information Technology Summit, (Confluence) 2014.

[4] Malladihalli S Bhuvan, Vinay D Rao, Siddharth Jain, Ashwin T S, Ram Mohana Reddy Guddeti, "Semantic Sentiment Analysis Using Context Specific Grammar" International Conference on Computing, Communication and Automation (ICCCA2015).

[5] Bing Liu, Minqing Hu, "Mining Opinion Features in Customer Reviews".

[6] Ankita Srivastava, Dr. M. P. Singh, Prabhat Kumar, "Supervised Semantic Analysis of Product reviews using weighted k-NN Classifier" 11th International Conference on Information Technology: New Generations, 2014.

[7] Soundarya.V, Siddareddy Sowmya Rupa, Sristi Khanna, G.Swathi, Dr.D.Manjula,"Extracting Business Intelligence from online product reviews "International Journal on Soft Computing (IJSC) Vol.4, No.3, August 2013.

[8] Hui Li, Shu Zhang, Xia Wang, "A personalization Recommendation Algorithm for E-commerce", Journal of Software, Vol. 8, No. 1, January 2013.

[9] Fung G., and Olvi L.M., "Incremental support vector machine classification." Proceedings of the Second SIAM International Conference on Data Mining, Arlington, Virginia. 2002.

[10] An approach towards comprehensive sentimental data analysis and opinion mining, 2014 IEEE.

[11] Sentiment analysis algorithms and applications: A survey, Walaa Medhat, Ahmed Hassan, Hoda Korashy.