International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 16
June 2018

Performance Comparison of Different Classifier Models In Sentiment Analysis

[1]G.Sailaja , [2]K.V. R.V.Prasanth Kumar , [3]Ch.Venkata Nagendra Babu , [4]A.Sai Phanindra
[1,2,3,4]Department of IT, Lakireddy Bali Reddy College of Engineering, Mylavram-A.P, India.
[1]sailajagade14@gmail.com,[2]kambhampati007@gmail.com, [3]venkatanagendrababumagee@gmail.com ,[4]saiphanindra.akula@gmail.com

Abstract

The goal of this paper is to analyse the performance of the different machine learning algorithms for data mining . In this study, 20 Machine Learning models were benchmarked for their accuracy and speed performance on different hardware architectures. These are applied when applied to 2 multinomial datasets differing broadly in size and complexity. Therefore, our study performs a benchmarking of different classification algorithms highlighting the adequacy and efficiency of different classifiers

Keywords

1. Introduction

Data mining is the process to pull out patterns from large datasets by joining methods from statistics and artificial intelligence with database management. It is an upcoming field in today world in much discipline. It has been accepted as technology growth and the need for efficient data analysis is required. The plan of data mining is not to give tight rules by analysing the data set, it is used to guess with some certainty while only analysing a small set of the data. Over the last years, the machine learning community has become increasingly aware of the need for statistical validation of the published results. This can be attributed to the maturity of the area, the increasing number of real-world applications and the availability of open machine learning frameworks that make it easy to develop new algorithms or modify the existing, and compare them among themselves.

The advent of Web 2.0 has led to an increase in the amount of sentimental content available in the Web. Such content is often found in social media web sites in the form of movie or product reviews, user comments, testimonials, messages in discussion forums etc. Timely discovery of the sentimental or opinionated web content has a number of advantages, the most important of all being monetization. Understanding of the sentiments of human masses towards different entities and products enables better services for contextual advertisements, recommendation systems and analysis of market trends. Amazon is one of the largest online vendors in the World. People often gaze over the products and reviews of the product before buying the product on amazon itself. But the reviews on amazon are not necessarily of products but a mixture of product of product review and service review. The buyer is misled as the overall sentiment (rating classification) that amazon gives is a collective one and there is no bifurcation between a service review and product review. The proposed model satisfactorily segregates service and product review, in addition to this it also classifies the review as feature review if the user talks about some particular product feature. A featured review is nothing but a product review, our model also gives sentiment of the text about the product feature. For example, if the user writes in his review, "the camera for this phone is very good.", then we also classify camera feature as positive. This paper develops an automated system which extracts the sentiments from the online posts from twitter. Our system shows sentiment identification, which expresses opinion associated with each entity. Also it consists of scoring phase, which assigns scores to each entity, on which the tweets are classified. Finally we have leveraged Naive Bayes, Support vector machine and other classifiers algorithms to do the sentiment analysis on this myriad of data.

2. Topic Extraction and Sentiment Analysis

Before WWW we were lacking a collection of opinion data, in an individual needs to make a decision, he/she typically asks for opinions from friends and families. When an organization needs to find opinions of the general public about its products and services, it conducted surveys and focused groups. But after the growth of Web, especially with the drastic growth of the user generated content on the Web, the world has changed and so has the methods of gaining ones opinion. One can post reviews of products at merchant sites and express views on almost anything in Internet forums, discussion groups, and blogs, which are collectively called the user generated content. Extracting features from user opinion information is an emerging task.

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 16
June 2018

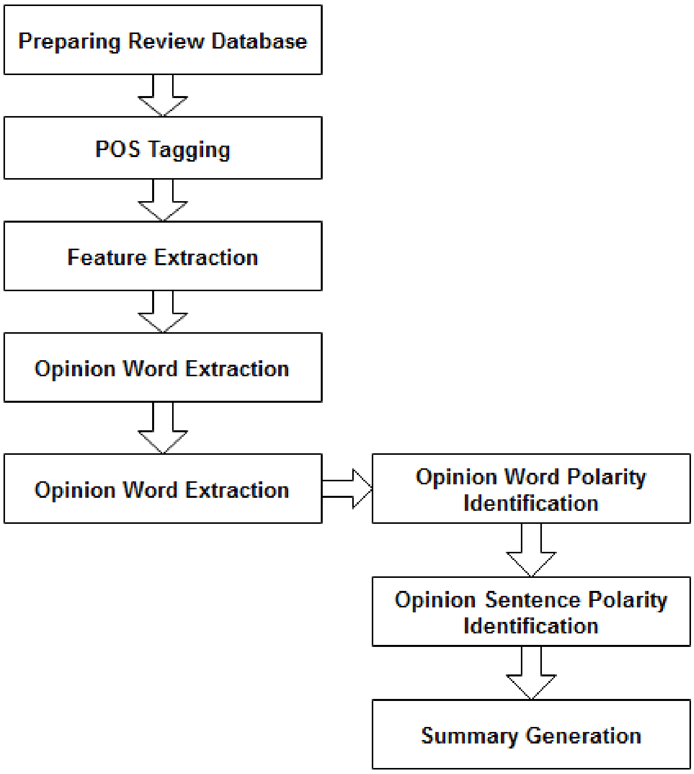

Fig 1. Basic Step of Feature Extraction

In Fig 1, a generic model of feature extraction from opinion information is shown, firstly the information database is created, next POS tagging is done on the review, next the features are extracted using grammar rules such as adjective + noun or so on, as nouns are features and adjectives are sentiment words. Next Opinion words are extracted followed by its polarity identification. Some models also calculate sentence polarity for accuracy. Lastly the results are combined to obtain a summary. Many algorithms can be used in opinion mining such as Naïve Bayes Classification, Probabilistic Machine Learning approach to classify the reviews as positive or negative, have been used to get the sentiment of opinions of different domains such as movie, Amazon reviews of products. For the purpose of this work our focus is on mining topics from user-generated product reviews and assigning sentiment to these topics on a per review basis.

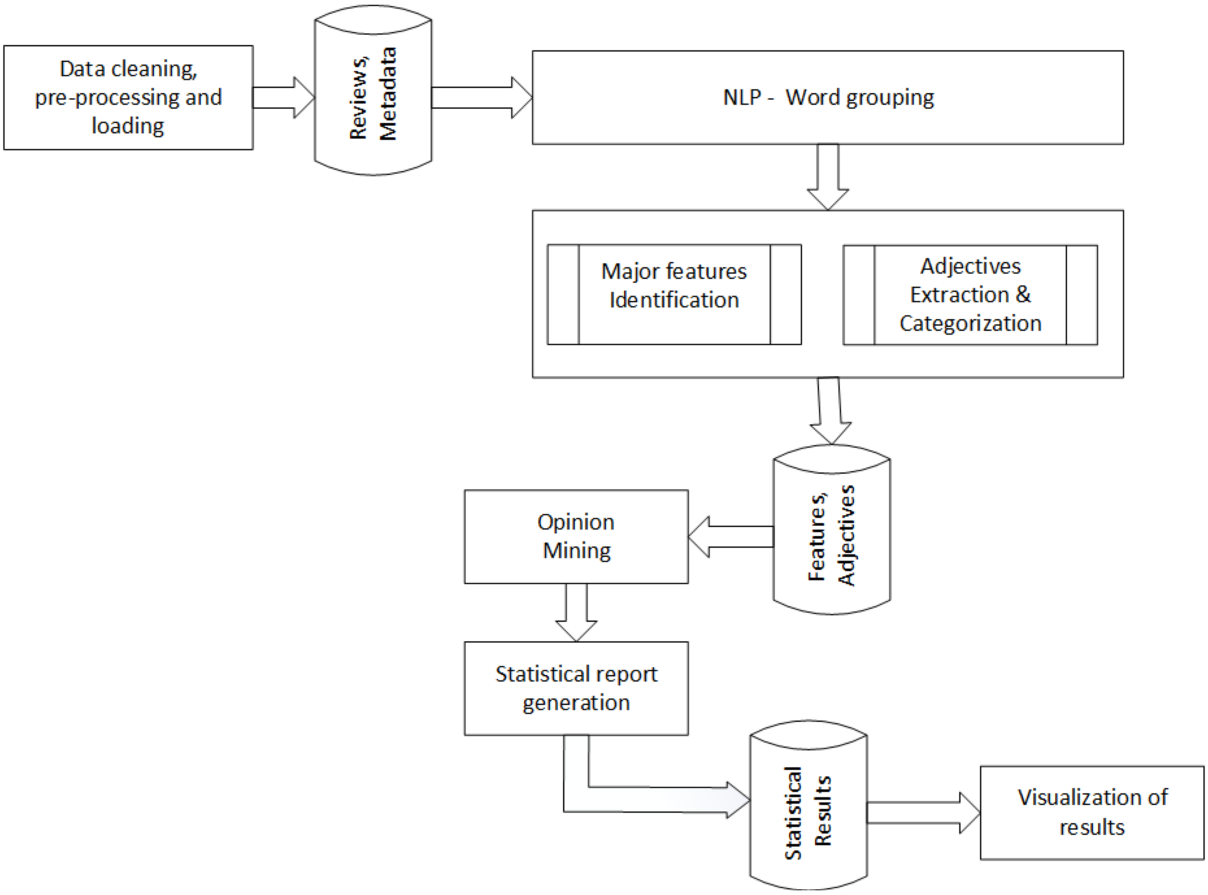

Fig 2. System architecture for extracting topics and assigning sentiment for user generated reviews

- **Data Pre-processing & Loading**

The available data is in JSON format files. The data pre-processing and loading into database using R

- **NLP- Word Grouping**

The review text undergoes Natural Language Processing to identify the noun phrases, verb and adjectives.

- **Major Features Identification**

The major features of a product that are mentioned by most of the reviewers are identified and stored in database.

- **Adjectives identification and categorization**

The adjectives mentioned about the features are identified and categorized as positive or negative and stored in database.

- **Opinion Mining**

Identifying whether a feedback gives positive or negative opinion on a feature of the product.

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 16
June 2018

- **Statistical Report generation**

Statistical data on the number of negative reviews for each feature of a product is saved.

- **Visualization of results**

The generated results are displayed in a user interactive portal.

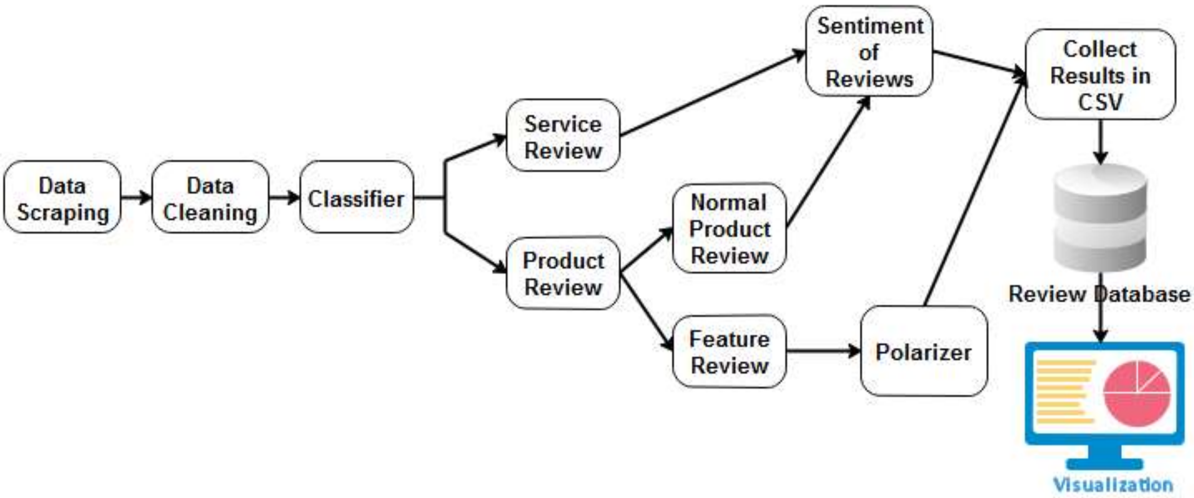

Fig 3. Detailed Block Diagram of System

- **Data Scraping**

Crawl the Amazon review url to extract all required details from it. We need to take care of the text so as to satisfy the required format, for e.g.
 tags have a special meaning to the browser i.e. break read or next line, we need to explicitly convert each
 tag to spaces or else the crawling result will be improper. When working with online reviews there is always a question in our mind, how can I trust the review. This is not a problem with Amazon reviews, Amazon reviewers can up vote or down vote a review, this collectively is available as helpful count. We have taken a special care in extracting the data from web pages smallest necessary data is extracted for processing. The following is the list of items that we have extracted: Review of Title, Helpful Count, User Review and Date of Review. Caution: Websites uses utf-8 character set for encoding characters, but sometimes this encoding can give errors during web scraping as scraping involves matching strings and patterns. Solution to this is simply enforce the string to be coded in utf-8 format.

- **Data Cleaning and Processing**

The data extracted need to be cleaned so that we get proper text review on which analysis can be performed. Cleaning of crawled data is done by removal of all special characters (such as: “:/.,’#$*^&-)

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 16
June 2018

in order to retrieve best results. After cleaning the crawled content copy it into a csv file. The next step is processing the cleaned data, firstly review is classified as service, feature or product review. If the review is a feature review then feature extraction is done using POS Tagging and grammar rule all stated below. After feature extraction the feature opinion polarization is obtained.

- All processed output is stored in one CSV file for further use.
- The file is then loaded into the database for use in visualization and summarization.
- Finally the summarization of sentiments is generated as charts and displayed to the user as an attractive dashboard

Algorithm for Feature Extraction

Step 1: Crawl the Amazon review url to extract all required details from it. Special care for required format of information must be taken, example
 tags have a special meaning to the browser i.e. break read or next line, we need to explicitly convert each
 tag to spaces or else the crawling result will be improper.

Step 2: Cleaning the crawled data. Removal of all special characters (such as : ":/.,'#$*^&-) must be done in order to retrieve best results. This also saves our review processing time. Put the crawled content into a CSV file.

Step 3: Read the CSV file for processing, for each review do the following:

Step 3.1: Perform a service review test where the review is tested for occurrence of service words, i.e. if the review length is shorter than 15 words and service words are found in the review the review is classified as service review else if the length of review is greater than 15 then more than 2 service must occur in the review for it to be a service review.

Step 3.2: If the review fails for the service test then it is tested for features of a product (such as camera, microphone and battery) if these exist then we classify the review as a feature review.

Step 3.2.1: For each feature we extract its sentiment from the review using POS tagging and ruled based extraction (using regular expressions).

Step 3.2.2: Each phrase of sentiment extracted above is then sent to polarizer that return 1 if the sentiment is positive else -1 which means the sentiment is negative.

Step 3.3: If the review fails the feature test also, then the review is classified as a product review.

Step 4: A new final CSV is generated with the classification and sentiment of the feature phrases.

Step 5: This CSV is then loaded into the database for creating the visualizations by querying data from the database.

3. Sentiment Polarity Methodology

The service and the product review's polarity is the rating the user provides for that review. The Good Reviews are those with rating 5 stars and 4 stars, Average Reviews are those with rating 3 stars and Bad Reviews are those with rating 2 stars and 1 star. Finally, when a feature sentiment is extracted the sentiment phrase is sent to a polarizer method, this method basically returns +1 if the phrase is a positive sentiment else -1 if the phrase is a negative sentiment. Firstly, the phrases are tested for indirect opinions such as "Battery no better than iPhone 4s", the test phrase is tested for certain pre-defined phrases that were found during manual analysis of reviews. Next if the phrase test fails, the review is tested for the word "not" if the word not exists then everything after not is polarized meaning every word after not is tested for whether it is a positive word or a negative word and consecutive words polarity are added and finally negated, for example "Camera is not good" this phrase is classified as negative as the word "good" is negated by the word "not". Lastly if "phrase" and "not" test fail the test phrase is broken down into words and polarity of each word is found from a dictionary of sentiment words bifurcated as good and bad words and collective polarity is considered i.e. if the sum is below 0 the outcome is negative (-1) else outcome is positive (+1).

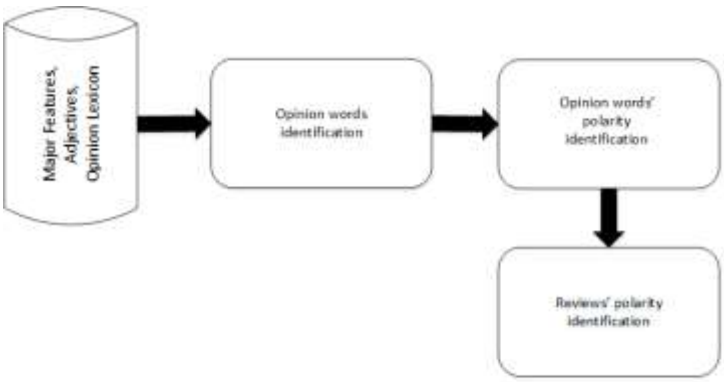

Fig 4. Low level functions of Opinion Mining

4. Rules for Feature Extraction

The following are some rules that our model uses to extract feature and its sentiment:

- Adjective + Noun
- Noun +Adjective
- Adverb + Noun
- Adverb +Adjective + Noun
- Noun + Adverb + Verb
- Noun + Verb
- Noun + Verb
- Noun +Verb + Noun
- Noun +Determiner + Adjective
- Noun + Verb + Adverb
- Noun + (verb or Adjective or Adverb)

5. Datasets & Setup

The review data for this experiment was extracted from Amazon.com during December 2017; 51,837 reviews from 1,384 unique products. We focused on 4 product categories—Digital Cameras (DC), GPS Devices, Laptops, Tablets — and labelled them as helpful or unhelpful, depending on whether their helpfulness score was above 0.75 or not, as described . For the purpose of this experiment, all reviews included at least 5 helpfulness scores (to provide a reliable ground-truth) and the helpful and unhelpful sets were sampled so as to contain approximately the same number of reviews. Table 1 presents a summary of these data, per product type, including the average helpfulness scores across all reviews, and separately for helpful and unhelpful reviews. Each review was processed to extract the classification features are described. Here we are particularly interested in understanding the classification performance of different categories of features. In this case we consider 8 different categories, AGE, RAT, SIZE, TOP, SENT-1, SENT-2, READ, CNT. Note, we have split the sentiment features (SENT) into two groups SENT-1 and SENT-2.

6. Results

The results are presented in Figures 5-8. In Figures we show the AUC performance for each classification algorithm (RF, SVM, NB) separately; each graph plots the AUC of one algorithm for the 8 different categories of classification features for each of the four different product categories (DC, GPS, Laptop, and Tablet). Figure 8 provides a direct comparison of all classification algorithms (RF, SVM, NB); here we use a classifier using all features combined. AUC values in excess of 0.7 can be considered as useful from a classification performance viewpoint. Overall we can see that RF tends to produce better classification performance across the various feature groups and product categories. Classification performance tends to be poorer for the GPS dataset compared to Laptop, Tablet, and DC. We know from previous research that ratings information proves to be particularly useful when it comes to evaluating review helpfulness. It is perhaps no surprise therefore to see that our ratings-based features perform well, often achieving an AUC > 0.7 on their own; for example in Figure 5 we see an AUC of approximately 0.75 for the Laptop and Tablet datasets, compared to between 0.65 and 0.69 for GPS and DC, respectively. Other 'traditional' feature groups (AGE, SIZE, READ, and CNT) rarely manage to achieve AUC scores > 0.7 across the product categories. We can see strong performance from the new topic and sentiment feature-sets proposed in this work. The SENT-2 features consistently and significantly outperform all others, with AUC scores in excess of 0.7 for all three algorithms and across all four product categories; indeed in some cases the SENT-2 features deliver AUC greater than 0.8 for DC, Laptop and Tablet products; see Figure 5. The SENT-2 feature group benefits from a combination of sentiment and ratings based features but a similar observation can be made for the sentiment-only features of SENT-1, which also achieve AUC greater than 0.7 for almost classification algorithms and product categories. Likewise, the topical features (TOP) also deliver a strong performance with AUC > 0.7 for all product categories except for GPS.

6. Conclusion

User-generated reviews are now an important source of knowledge for consumers and are known to play an active role in decision making in many domains. In this paper we have described techniques for mining topical and sentiment information from user-generated product reviews as the basis for a review quality classification system. We have demonstrated that these topical and sentiment features help to improve classification performance above and beyond that which is possible using more conventional feature extraction techniques. We have further described a possible application of this classification approach and evaluated its ability to make high quality review recommendations in practice.

International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 16
June 2018

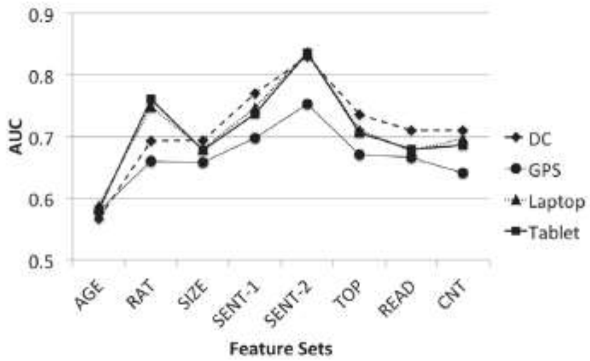

Fig 5 Performance using Naive Bayes Classifier

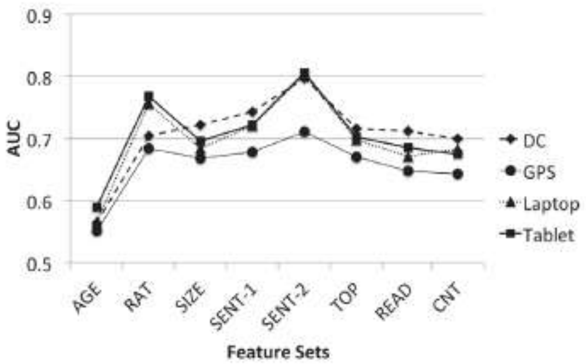

Fig 6. Performance using SVM classifier

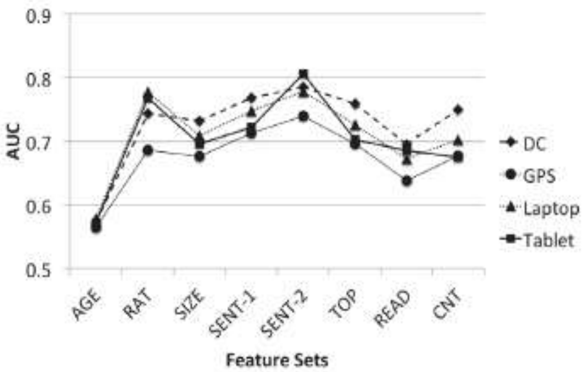

Fig 7. Performance using Random Forest Classifier

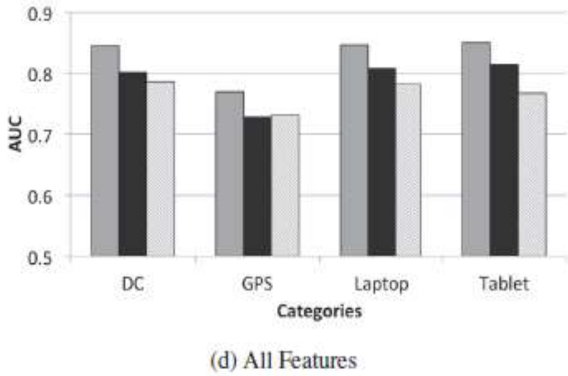

Fig 8. Performance for all the classifier

REFERENCES

[1] J. Ashok Kumar and S. Abirami, "An Experimental Study of Feature Extraction Techniques In Opinion Mining," International Journal on Soft Computing, Artificial Intelligence and Applications (IJSCAI), Vol.4, No.1, February 2015.

[2] GurneetKaur and AbhinashSingla, “Sentimental Analysis of Flipkart reviews using Naïve Bayes and Decision Tree algorithm,” International Journal Of Advanced Research in Computer Engineering & Technology (IJARCET), Vol. 5, ISSN: 2278- 1323, January 2016

[3] Mohamed Ibrahim Abouelhoda , Stefan Kurta and Enno Ohlebusch,”Replacing suffix trees with enhanced suffix arrays”, 1570-8667, 2004.

[4] Aashutosh Bhatt, Ankit Patel, Harsh Chheda and Kiran Gawande, "Amazon Review Classification and Sentiment Analysis", International Journal of Computer Science and Information Technologies, Vol. 6, ISSN:0975-9646, 2015.

[5] Jeevanandam Jotheeswaran and S. Koteeswaran,“Feature Selection using Random Forest method for Sentiment Analysis,” Indian Journal of Science and Technology, vol.9. ISSN : 0974-5645, Jan. 2016.

[6] Lisette García-Moya, Henry Anaya-Sánchez, and Rafael Berlanga-Llavori, “Retrieving Product Features and
Opinions from Customer Reviews”, IEEE Intelligent Systems, vol. 28, pp. 1541-1672, May/Jun 2013.

[7] Tina R. Patil and Mrs. S. S. Sherekar, “Performance Analysis of Naive Bayes and J48 Classification Algorithm for Data Classification,”

International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 16
June 2018

International Journal Of Computer Science And Applications, vol. 6, ISSN: 0974-1011, Apr 2013.

[8] Dim En Nyaung, Thin Lai Lai Thein, "Feature Based Summarizing From Customer Reviews", International Journal Of Scientific Engineering and Technology Research, vol.03, ISSN 2319-8885, Dec-2014.

[9] Bing Liu, Sentiment analysis and opinion mining, Morgan & Claypool Publishers, 2012.

[10] R Krishna, Kavitha S, Yamini S,Rekha A,"Analysis of Various Opinion Mining Algorithms Gayathri," International Journal of Computer Trends and Technology, vol. 22 , Number 2, ISSN: 2231-2803 ,April 2015.