Website Evaluation Using Opinion Mining

Aditya Bhatnagar[1], Maaz Nomani[2], Ketan Pujare[3], Dr. Sunil Wankhade[4].

Bachelor of Engineering, Dept. of Information Technology, R.G.I.T., Mumbai, Maharashtra, India[1]

Bachelor of Engineering, Dept. of Information Technology, R.G.I.T., Mumbai, Maharashtra, India[2]

Bachelor of Engineering, Dept. of Information Technology, R.G.I.T., Mumbai, Maharashtra, India[3]

Head and Professor, Dept. of Information Technology, R.G.I.T., Mumbai, Maharashtra, India[4]

Abstract: With the advancement in technology, online shopping has taken a new toll. Customers can access the shopping websites from anywhere with the help of the internet. The websites display all the products for the customer to choose from and the customer can post any review or their opinion on the service of the website. The main aim of this project is to provide users with a rating generated according to the shopping website's performance and standards. There are numerous shopping websites to visit but the legitimacy of all the websites is questionable. Fraud can occur on a regular basis from illegitimate websites. The users may not get the service as per their expectations. There are many renowned shopping websites which do not match their credibility and in turn disappoints the customers or users. The proposed system evaluates the shopping websites based on the reviews or opinions of the experienced customers. The performance of the website is processed down to certain parameters. According to these parameters a certain rating of that particular website is generated, which helps the customer to decide which website to opt for.

Keywords: Natural language processing , Natural language toolkit(NLTK), decision making and linguistics.

_____ ********* _____

I. INTRODUCTION

The large number and the variety of existing Web-sites and their increasing role in social and economic contexts demand for flexible methods and tools to support Website quality evaluation in an efficient and efficacious way. As users, and even more as professionals, we have all experienced that for many websites navigation is all but satisfactory, needed information is missing, transaction forms are awkward, graphical design is shabby, analytics show low popularity and success, and so on. Every passing day the number of companies, organizations and individuals publishing their web sites is increasing. Considering all the information available on the web every individual should desire to find and access useful information. For example users want to learn about different shopping websites and what products and services they offer using the web. By the help of this information users may learn about the websites and in turn choose a website which is suitable according to their standards. The task of evaluating and improving the web sites can be intimidating considering the number of web sites available and the frequency of updates. As a result, automated support for web designers and web site owners become more important. Automated usability tools can help save time and money in design. User testing can improve consistency and quality of the web site. The customers wanting to buy the product would like to compare the products and the services provided by different websites before purchasing the product. The services and all the parameters related to the website can't be compared on the existing system . Therefore it makes it difficult for the customers to decide. Although the websites would be rated but comparing the websites would make it more efficient and can provide a clear to opt for the better product. Hence it was highly important to come up with the solution for the customers which would be the evaluation of websites based on the review of the customers and rate it according to the sentiments identified in the reviews.

II. LITERATURE REVIEW

The aim of Machine Learning is to develop an algorithm to increase the performance of the system using the given data or past experience. The Machine Learning provides a

solution to the classification problem that involves two steps: 1. Learning the model from a corpus of training data. 2. Classifying the unseen data based on the trained model. Sentiment analysis of natural language texts is an emerging field. Converting a part of text to a feature vector is an essential step in any data driven approach to Sentiment analysis . Lisette García-Moya, et all addresses the aspect-based summarization task by introducing a novel methodology for retrieving product features from a collection of free-text customer reviews about a product or service. Their proposal relies on a language modeling framework that combines a probabilistic model of opinion words and a stochastic mapping model between words to approximate a language model of products. Their work extends a preliminary approach introduced which addresses the modeling of a language of product features from customer reviews. They provide a formalized methodology for the retrieval of product features from the estimated language model of features.

Madhavi Kulkarni and Mayuri Lingayat proposed a technique which ranks efficiently the products by mining the genuine reviews of the product. System provides a method which allows only those users to write a review about a product who have purchased from the website. Other users are not allowed to give review. This reduced the wrong reviewing of product and customer get reliable product. They have proposed product ranking system which can take product information in query form and system provides the product matching with customer's requirements along with product ranking. Arti Buche, paper provides information about the data sources that opinion mining uses, machine learning and tasks of sentiment analysis for sentiment classification, text classification and the tools available for sentiment classification, and the performance evaluation. Red Opal tool helps users to recognize products which rely on their
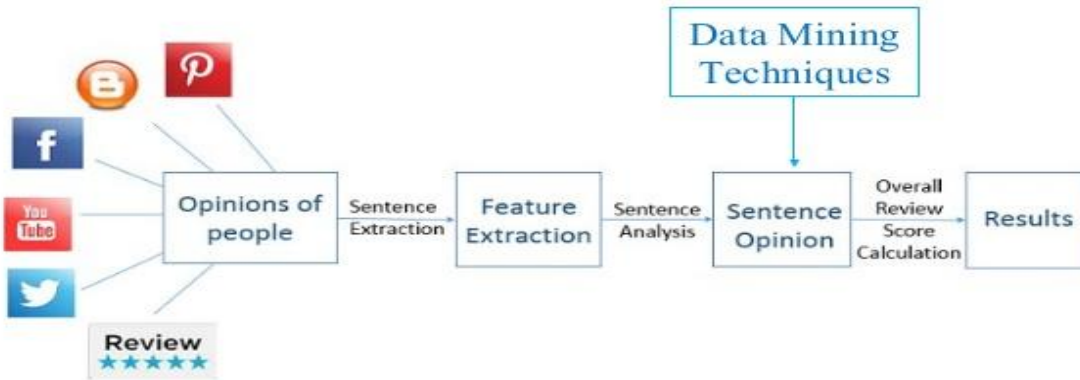
features. Opinions of products are given in graph format feature by feature. Review Seer tool is used for automation and aggregation of sites. Naïve Bayes classifier is used to assign score to the extracted features which are done on classified reviews.

Aashutosh Bhatt proposed a system that performs the classification of reviews given by customer which is then followed by finding sentiment for the reviews given. A rule based extraction of the feature for the given product is done. Their system visualizes the review's sentiment which is then presented in the form of a chart. Classification of reviews along with sentimental analysis increased the accuracy of the system which in turn provided accurate reviews to the user. The main aim of the system is to ensure fair results of sentiments and saving time, that user used to spend on reading long textual descriptions of reviews, by providing summarized results in charts. Asmita Dhokrat, et al. [9] gave review about the techniques and tools used in opinion mining. Paper shows the basic requirements of opinion mining to explore the present techniques used to developed a full fledge system. It highlights the opportunities or deployment and research of such systems. The available tools for building such applications are presented with merits and limitations. Farhan Hassan Khan proposed a new Twitter Opinion Mining framework to predict the polarity of words into positive or negative feelings in tweets. It improves the accuracy level of the classification. It is constructed using different stages.

III.   PROPOSED SYSTEM

Website evaluation is a software developed to evaluate or rate a website based on the sentiments or opinions of users given to those websites. It facilitates the customers with parameters covering all the aspects required by the user to choose a suitable website according to their requirement standards. This system will also protect users from fraudulent and theft related shopping websites.

Figure 1: Block diagram of Website evaluation using Opinion Mining

Opinions of the users are extracted from the APIs and the string is divided for the further feature extraction. The reviews of the users is extracted and performed sentiment analysis using the Natural language Toolkit(NLTK) which is an existing library of python.The sentences are broken down and thus performed analysis to give a rating about the website which will be displayed on our webpage.
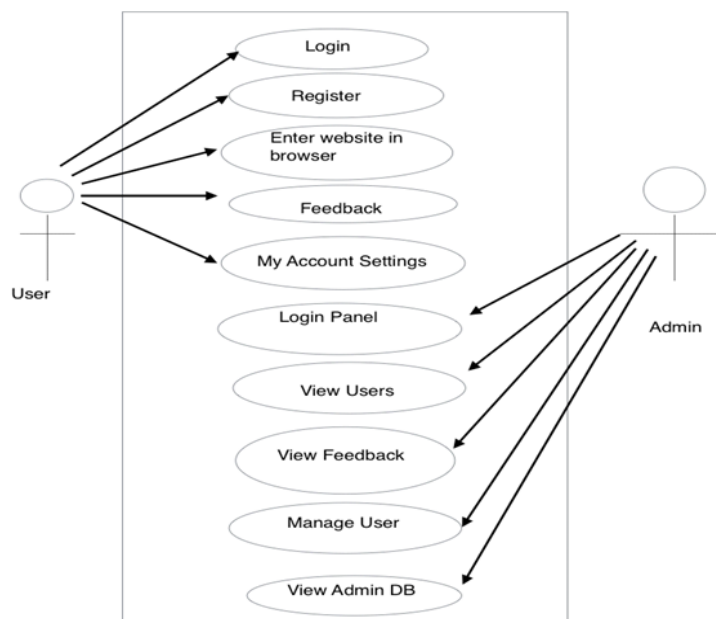
IV. ANALYSIS

Figure 2: Use Case Diagram of Website evaluation using Opinion Mining

In fig. 2, There are two actors; The admin and user. The user can login or register into our website depending on the registered account of the user. The user then enters the name of the website to acquire the rating generated based on

opinions. However the admin can perform activities like managing the login panel, view and manage users and manage admin database.
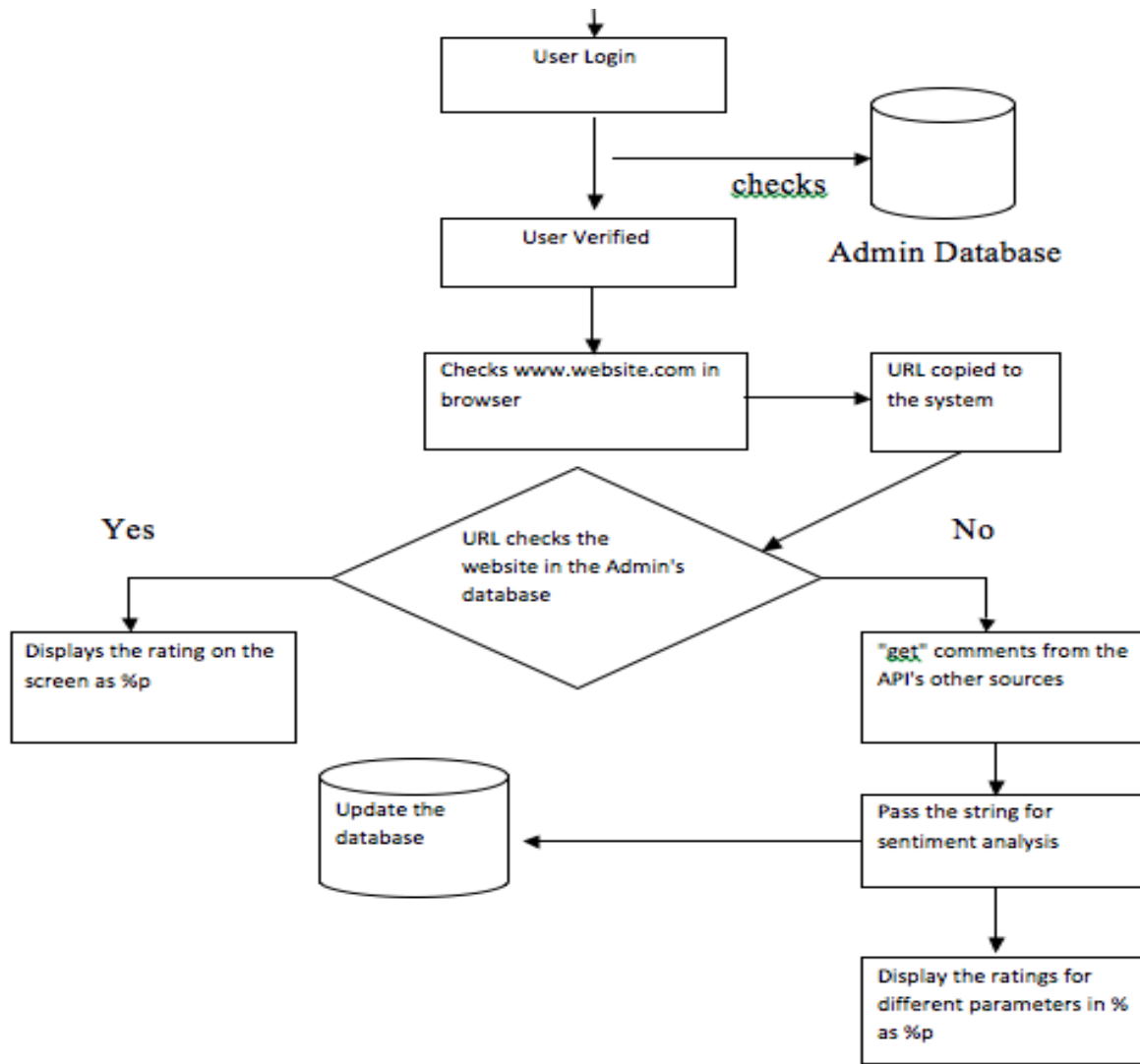

Figure 3: Activity Flowchart diagram of Website evaluation using Opinion Mining

In fig.3, the activities that are to be performed ; The user logs in the system and the account of the user is verified, the user then enters the name of the desired website in the input text area. The system then checks the website in the

DB, if available the system directly displays the result on the webpage . However if the website isn't available in the DB the data is extracted , analyzed & displays the result . The result is also updated in the DB.
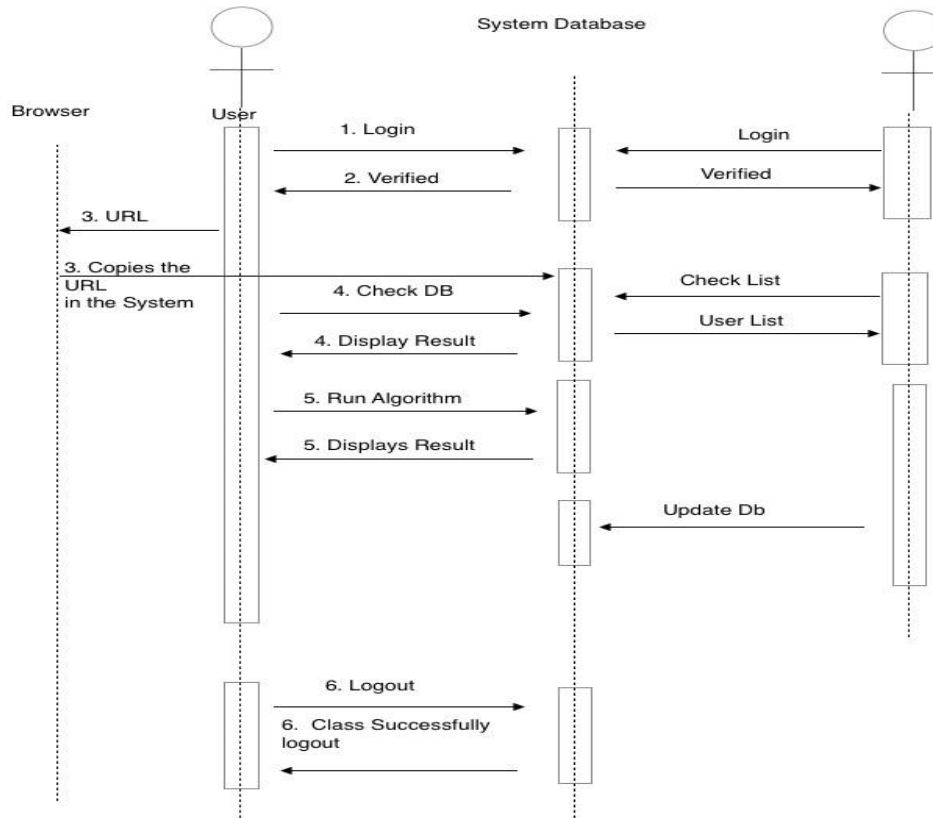
Figure 4: Sequence diagram of Website evaluation using Opinion Mining

In fig. 4, the sequence of the activities is shown in the diagram. The user logs in , it's authenticity is checked and verified and allows the user to login in the system. The user then enters the name of the website, it is copied from the browser to the backend where the algorithm runs, and the result is displayed on the webpage.
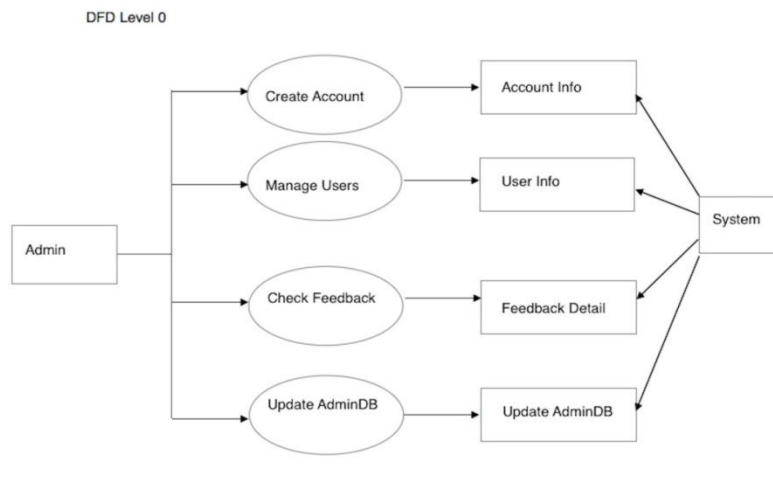
V. DESIGN

Figure 4: DFD Level 0 Diagram of Website evaluation using Opinion Mining

In fig. 4, the Data flow at level 0 is shown. It shows the functions performed by the admin and thus showing the connectivity between the system.
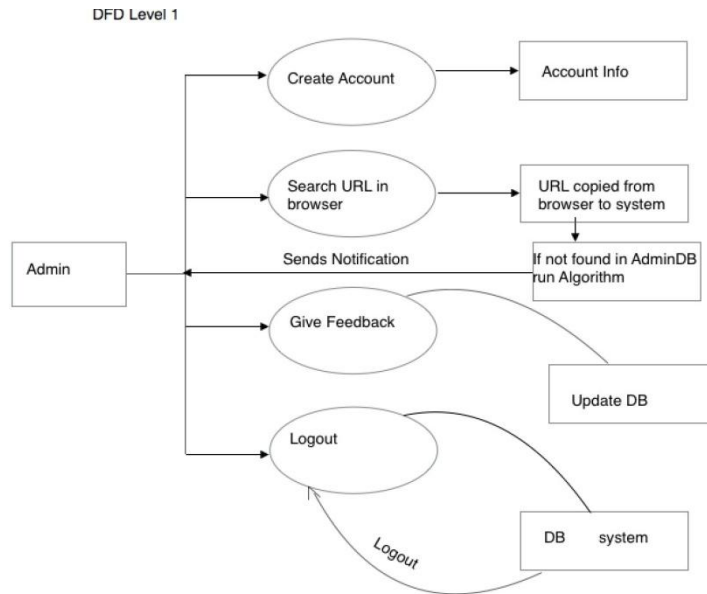

Figure 5: DFD Level 1 Diagram of Website evaluation using Opinion Mining

In fig. 5, the Data flow at level 1 is shown. The admin performs activities like entering the website in the input field and the website name is copied in the system. The system checks it the DB if not present runs algorithm, displays the result and updates the DB.
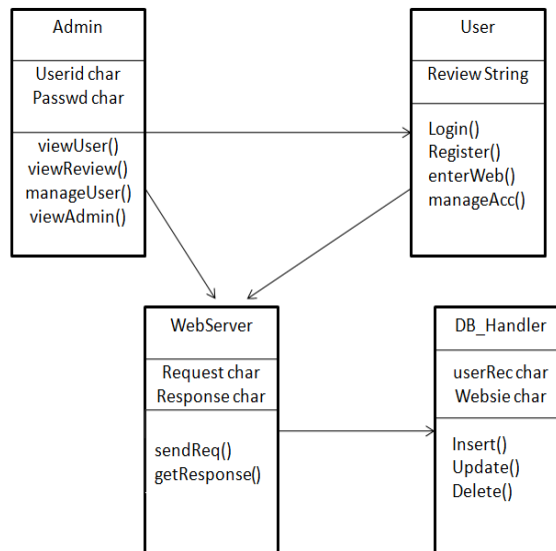

Figure 6: Class Diagram of Website evaluation using Opinion Mining

In Figure 6, the classes present are Admin, User, Webserver and DBhandler. Webserver is the center of this

project. Second row contains all the attributes related to that class. Third row in class diagram are methods or operations

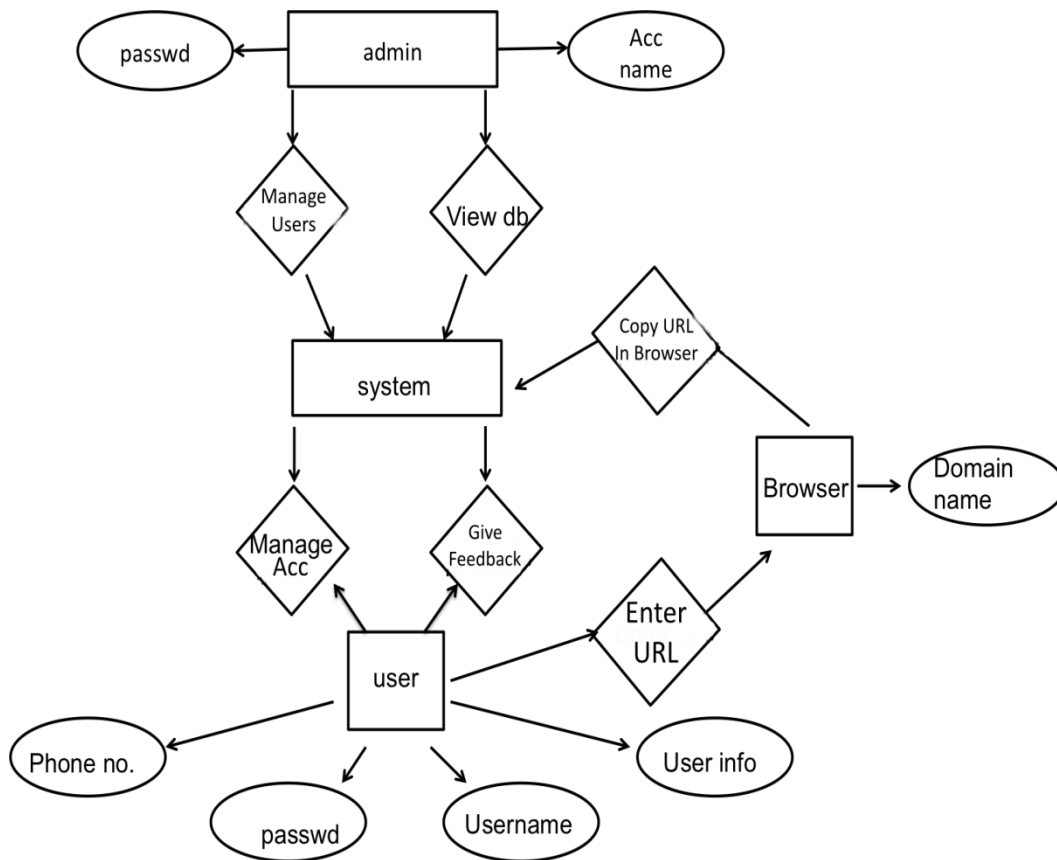that are to be performed regarding that particular class.

Figure 7: ER Diagram of Website evaluation using Opinion Mining

Different entities present in fig. 7 are admin , webserver, browser and user. The admin is related to the system with the help of DB, whereas the system is related to the user by

feedback and account management property. The user has attributes like username, password & phone no.

CONCLUSION

Our system gives the customer a better understanding about e-shopping websites. It prevents the customer from being a victim of fraud. It will give them a detailed survey about the service or customer feedback regarding those shopping websites. Our system could be a critic for the shopping sites hence driving them towards improvement on their drawbacks. .On the basis of data availability our system can be expanded to different sectors. It is also open to customer's views based on their frequent shopping experience with those websites.

V.   ACKNOWLEDGEMENT

VI.   REFERENCES

[1] Gautami Tripathi, S Naganna, "Opinion Mining: A Review", International Journal of Information & Computation Technology.

[2] P Haseena Rahmath, "Opinion Mining and Sentiment Analysis - Challenges and Applications", International Journal of Application or Innovation in Engineering & Management (IJAIEM), vol. 3, no. 5, May 2014.

[3] Nidhi R. Sharma, Vidya D. Chitre, "Opinion Mining Analysis and its Challenges", International Journal of Innovations & Advancement in Computer Science, vol. 3, no. 1, April 2014.

[4] S. Chandra Kala, C. Sindhu, "Opinion Mining And Sentiment Classification: A Survey", ICTACT Journal on Soft Computing, vol. 3, no. 1, October 2012.

[5] G. Vinodhini, R.M. Chandrasekaran, "Sentiment Analysis and Opinion Mining: A Survey", International Journal of Advanced   Research in Computer Science and Software Engineering, vol. 2, no. 6, June 2012.