# Sentiment Analysis from Customer Reviews – A Survey

**Bharti Devendra[1], Shiv Shakti Shrivastava[2]**
[1,2]Rabindranath Tagore University, Bhopal (M.P.) India.

**ABSTRACT**
*Today reviews of the customers on any product or services are highly recommended and appreciated by the online users to come up with a decision. Sentiment and opinions made by a person can influence the perceptions of the other. The extensive use of internet and the growth of web services among the people also created an enormous data in terms of the reviews towards certain services or products. That data can easily be accessed worldwide and assist an individual to createhis own viewpoint. The key objective of gathering opinions of the customers is to assess the quality and other parameters of products and services. This paper proposesa methodologyfor the analysis of sentiments by classifyingthe feedbacks. The proposed project also reviewed the existing systems, which projects the technique of sentiment analysis.*

***Keywords***-Sentiments, reviews, opinion mining, customers.

## I INTRODUCTION

These days, either the services or the products offered by any firm needs an opinion from their customers to enhance the quality of the product or the services provided by them. There are number of vendors available in the market to offer various online services. Reviews of the customers will not only assist those vendors to improve the quality of services provided by them but also help other customers to know about them and come up with a cumulative decision. There are several websites, different forums, blogs and social media application available, which offer the facility of reviewing the product or services provided by them. The feedback given by a particular customer can be randomly read by other customers and accordingly make opinions about that particular product or services. As the reviews about any particular product or the services rendered are higher in number, their opinions will vary too. This mayleada confused and impartial opinion towards that product or service. Opinion mining for sentiment analysis is an effective technique, which extracts the information by text processing and improves the decision making of a customer while using any online services. Customer's reviews are sometime biased, as their ratings and comments are not so coordinated. The proposed research focuses on the analysis of reviews made by the customers and uses the approach to analyze the sentiments. This will create a classification of reviews as positive, negative and neutral, which is known as sentiment polarity [1].

**(a) Classification of Existing Solutions** -Prior work done on sentiment analysis can be distinguished in terms of different factors like; used method, textual view, level of extraction of details from textual analysis, review and rating level etc.Technically, there are three different approaches widely used for sentiment analysis.

**(i) Machine Learning Approach :** Machine learning approach is used to analyze the sentiment by using supervised techniques which needs to get trained in different preferable domain and the approach uses various algorithms of learning. On the other hand statistical unsupervised learning do not require any prior training and can be dynamically used but at the stake of precision.

**(ii) Lexicon-Based Approach**: This kind of approach computes the polarity level of the sentiments from the reviews of the customers or the users using the semantic notation of words or sentences mentioned in any of the given opinions or reviews. Semantic Opinion is nothing but a calculation of outlook in terms of opinion in the mentioned text and the subject.

**(iii) Statistical and Rule-Based Approach:** This type of approach analyzes the opinion in terms of words mentioned in a text and distinguishes it depending on the number of positive words and negative words. There are different rules taken into consideration for this approach like dictionary polarity, negation words, booster words, idioms, emoticons, mixed opinions etc.

**(b) Rating Levels:**

**(i) Aspect Rating**

**(ii) Global Rating**

A method can be classifiedbased on two parameters. Computing the strength of the sentiments from various aspects of the services and the products provided by the merchant is considered as the aspect rating, whereas the technique used to rate services and products worldwide at the global level comes under the global rating where throughout the world people can review or read the reviews of certain products or the services. Many services prefer global rating of reviews, which takes mere the polarity of the opinions in terms of positive, or negative and processed on machine learning platform. A technique, which focuses on thorough categorization of ratings (like 4 stars out of 5), needs more linguistic methods, which include different structures. Figure No. 1shows the categorization of presented methods.

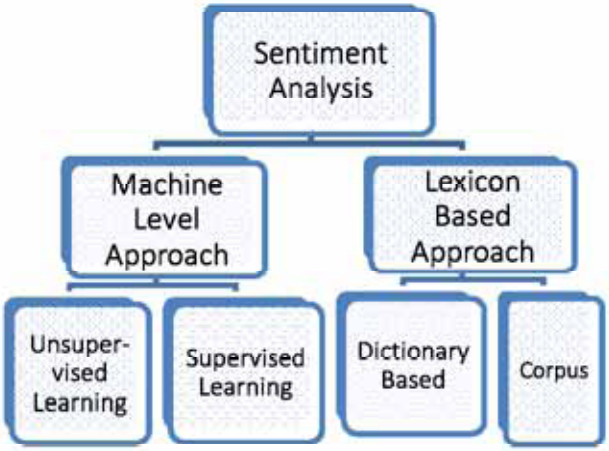


**Fig No. 1. Sentiment Classification**

(c) **Tools for Sentiment Analysis:** There are various tools available for the analysis of sentiments and study the opinions given by users through their ratings and reviews. Commonly used systems consist of different tools to analyze those reviews from different aspects and parameters and conclude it in the category of positive or negative reactions. The Table No. 1gives the categories of the tools available and their purpose in the system.

**Table No. 1.**
**Tools for Sentiment Analysis**

| TOOLS FOR SENTIMENT ANALYSIS | TECHNIQUES USED BY TOOLS |
|---|---|
| EMOTICONS | Emoticons contained in the Text |
| LIWC | Dictionary & Sentiment Classified Categories |
| Senti Strength | LIEC Dictionary with new features to strength & weak sentiments |
| Senti Word Net | Lexical Dictinary & Scores Obtained by Machine Learning aproaches |
| SenticNet | Natural Language Procesing Approach for inferring the polarity of semantic level |
| Happiness Index | Affectve Norms for English Words (ANEW) & Scores fpr evaluating happiness in the text |
| AFINN | Affective Norms for English Words (ANEW)but more focussed on the language used in microblogging platforms |
| PANAS-4 | Eleven sentiment psychometric scales |
| Sentiment140 | API that allows tweets classfying tweets t Polarity classes, positive, negative & neutral |
| NRC | Large sets of human provided words with their emotinal tags |
| EWGA | Entroy weighted genetic algorithm |
| FRN | Feature relation network considering syntatic n-gram relations |

## II LITERATURE SURVEY

### (a) Review on existing systems:

**Shraddha Mehtaet. al. [2]**
This paper proposed a method to analyze the rate of success of a movie. For this, an algorithm of ID3 classification was used to create the conclusive rules. The produced conclusions were utilized to calculate the rate of success of a movie. The obtained result was further utilized by the researchers of that field to recognize those movies, which may underperform at the box office. But while using the ID3 algorithm, at a time, single test was being executed which increased the decision instance. Even small testing resulted in an over classified set of data and the classification of constant data increased the expenses of calculation.

**Jingbo Zhu et. al. [3]**
Traditionally, to obtain the reviews of the customer about the centered product or services, a series of questions were generated to study the satisfaction rate of the customers who were served. For this, opinion

polling was done. The proposed paper uses the method of aspect based bootstrapping to study the opinions of the customers which neither needs any review in textual form nor any customer to reply the question generated to check the service quality. The proposed algorithm studies aspects to identify and creates segmentation of multi aspect sentences. In the proposed survey, system achieved 75.5% of accuracy during polling. However, the proposed technique does not offer guarantee of the taken samples and fundamentally enhances the load of calculation.

**Xiaohui Yu et. al. [4]**

The proposed technique in the paper is concerned about the data obtained by developing algorithms and methods, which is useful to gather the information captured from the reviews of the customers. The proposed model has the capability to forecast the upcoming sales of any product, which leads to the steps taken by the stakeholder. A case study is organized for the domain of movies and studies the issues for the prediction of the performance of sales. The analyzed result obtained observes that the sentiments shown in the reviews and their quality have a considerable impact on the upcoming performance of sales of product. While considering sentiment factor, Sentiment PLSA(S-PLSA) is developed and to enhance the accuracy of prediction. Here, the system takes quality factor as a parameter. Though the completely proposed system only uses S-PLSA, it can be implemented in many other scenarios to enhance the performance. For future work, they aspire to discover their role in clustering and differentiation, which is based on the sentiments.

**Guoshuai Zhaoet. al. [5]**

The paper proposed a method, which emphasizes on the user's reviewing manners by taking four parameters related to social network, which are of personal interest of user, interpersonal interest similarity, interpersonal rating behavior similarity and interpersonal rating behavior diffusion [5]. The proposed method of scheduled rating shows the reviewing and rating behavior of the users. The author surveyed the completely social circle of the users and differentiates them into three categories direct, indirect and mutual friends to analyze the rating behavior diffusion [5]. All the resultant factors are taken into consideration to enhance the precision and implementation of the predictions. Taking these factors to explore the rating behavior of the users will also increase the junk information in the database and increased parameter to analyze the behavior may end up with the biased results.

**Muhammad Taimoor Khan el. al. [6]**

Thorough study and deep analysis of the issues by using ML (machine learning) based technique shows that SA (Sentiment Analysis) and NLP (Natural Language Processing) have some drawbacks, whichcannot control the used method. As in the method of NLP, sentiment analysis faces issues like co-reference resolution, negation handling, and word sense disambiguation etc.[6] includes more problems because of the variations in the considering factors. It has been observed that the method of SA (Sentiment analysis) is nothing but highly constrained NLP problem because the trained system is not efficient enough to identify with the semantics of every word. Complex network analysis is one of the frequently used method for various issues and able to show constructive patterns in subjective text [6]. Machine learning classifies in supervised and unsupervised techniques. Supervised techniques need to be trained in some applicable platform or domain and are more precise. Unsupervised techniques are already trained and are proficient in dynamic environment but at the cost of precision [6].

**Silvana Aciar et. al.[7]**

Examining the features of a movie by using the technique of machine learning is comparatively unexplored technique to predict the success rate. By taking into consideration, the collected data turns into useful information not only in the interest of the producers and financiers but also the service providers and the viewers. Mostly the existing work emphasizes user specific preference and examines the reviews of the movies [7]. Due to the relevance of huge number of factors and attributes with the success of a movie, various works have been already done in this field. The classification model developed in the proposed paper shows effective results in prediction of the ratings of movies before their release based on movie information database (IMDb). The proposed method also approaches to mine unstructured data from movie [7].

## III PROPOSED METHODOLOGY

This study aims at

(i) Exploring how sentiment analysis and opinion mining strategies can be applied to textual reviews to extract semantic orientations on products, events; and to demonstrate how these opinions can be used for trend detection and knowledge discovery.

The study proposes following objectives:

(ii) To conduct a literature survey on sentiment analysis strategies and identify suitable domain for analysis domain (like iphone6, laptop).

(iii) To develop language analysis strategies that are applicable to pre-process text content from social media.

(iv) To compare the proposed model with existing ones and to evaluate the usefulness of processed knowledge.

(v) Customers are mainly nontechnical persons. Therefore, the method aims to envision results in legible format like graphs, charts or summary.

## IV CONCLUSION

Effective sentiment analysis algorithm/ modules for the Aspect based Sentiment Analysis and visualization of textual feedback will be developed.

## REFERENCES


[1] The Analysis andPredictionofCustomerReview Rating Using OpinionMining proposed by Wararat Songpan, DepartmentofComputer Science, Faculty ofScience, Khon KaenUniversity, Khon Kaen, Thailand, 978-1-5090-5756-6/17/$31.00 ©2017 IEEE SERA 2017, June 7-9,2017, London, UK

[2] Shraddha Mehta1, Hitarthi Bhatt and Prof. Darshana Desai, "A Compendium for Prediction of Success of a Movie Based Upon Different Factors", IJARCCE ISSN 2278-1021, Vol. 4, Issue 12, December 2015.

[3] Jingbo Zhu, Member, IEEE, Huizhen Wang, Muhua Zhu, Benjamin K. Tsou, Member, IEEE, and Matthew Ma, "Aspect-Based Opinion Polling from Customer Reviews", IEEE Transactions On Affective Computing, Vol. 2, No. 1, January-March 2011

[4] Xiaohui Yu, Yang Liu, Jimmy Xiangji Huang, and Aijun An, "Mining Online Reviews for Predicting Sales Performance: A Case Study in the Movie Domain", IEEE Transactions On Knowledge And Data Engineering, Vol. 24, No. 4, April 2012.

[5] Guoshuai Zhao, Xueming Qian, and Xing Xie, "User-Service Rating Prediction by Exploring Social Users' Rating Behaviors" IEEE Transactions On Multimedia, Vol. 18, No. 3, March 2016.

[6] Muhammad Taimoor Khan, Mehr Durrani, Armughan Ali, Irum Inayat, Shehzad Khalid and Kamran Habib Khan, " Sentiment Analysis And The Complex Natural Language", Complex Adapt Syst Model, DOI 10.1186/s40294-016-0016-9, 2016.

[7] Silvana Aciar, Debbie Zhang, Simeon Simoff, and John Debenham, "Informed Recommender: Basing Recommendations on Consumer Product Reviews", The IEEE Computer Society 1541-1672/07 39, 2007.

[8] Xiao Sun, Chongyuan Sun, Fuji Ren, Fang Tian, Kunxia Wang, "Emotional Element Detection and Tendency Judgment Based on Mixed Model with Deep Features", 978-1-5090-0806-3/16, IEEE ICIS, June 2016

[9] Zhongchen Miao, Junchi Yan, Kai Chen, Xiaokang Yang, Hongyuan Zha, and Wenjun Zhang, "Joint Prediction of Rating and Popularity for Cold-Start Item by Sentinel User Selection", Digital Object Identifier 10.1109/ ACCESS.2016. 2633282, 2016

[10] Xiaojiang Lei, Xueming Qian and Guoshuai Zhao, " Rating Prediction Based on Social Sentiment From Textual Reviews IEEE Transactions On Multimedia, Vol. 18, No. 9, September 2016

[11] G. Sneka, C. T. Vidhya, "Algorithms for Opinion Mining and Sentiment Analysis: An Overview", International Journal of Advanced Research in Computer Science and Software Engineering Volume 6, Issue 2, ISSN: 2277 128X, February 2016

[12] Filipe N Ribeiro, Matheus Araujo, Pollyanna Goncalves, Marcos Andre Goncalves and FabricioBenevenuto, "SentiBench- A Benchmark Comparison of State-Of-The-Practice Sentiment Analysis Methods", EPJ Data Science, DOI 10.1140/epjds/s13688-016-0085-1, 2016.

[13] Yuanlin Chen, Yueting Chai, Yi Liu, and Yang Xu, "Analysis of Review Helpfulness Based on Consumer Perspective", Tsinghua Science And Technology, pp 293-305 Volume 20, Number 3, June 2015.

[14] Abhishek Kaushik, Anchal Kaushik, Sudhanshu Naithani, "A Study on Sentiment Analysis: Methods and Tools",IJSR, Volume 4 Issue 12, December 2015

[15] Rahul M. Sharma, Shital S. Barkul, et al., "Online Movie Review System" Multidisciplinary Journal of Research in Engineering and Technology, Volume 2, Issue 3, Pg.547-552, 2015

[16] Andrei Oghina, Mathias Breuss, Manos Tsagkias, and Maarten de Rijke, "Predicting IMDB Movie Ratings Using Social Media" ISSN 2278-1021, IJARCCE DOI 10.17148, 2015.

[17] Kewen Wu, Zeinab Noorian, Julita Vassileva and Ifeoma Adaji, "How buyers perceive the credibility of advisors in online marketplace: review balance, review count and misattribution", Journal of Trust Management, DOI 10.1186/s40493-015-0013-5, 2015

[18] Ananta Arora, Chinmay Patil, Stevina Correia, "Opinion Mining: An Overview", Vol. 4, Issue 11, International Journal of Advanced Research in Computer and Communication Engineering, DOI 10.17148/IJARCCE.2015.41122 94, November 2015.

[19] Syed Akib Anwar Hridoy, M. Tahmid Ekram, Mohammad Samiul Islam, Faysal Ahmed and Rashedur M. Rahman, "Localized Twitter Opinion Mining Using Sentiment Analysis", Decision Analysis:A SpringerOpen Journal, DOI 10.1186/s40165-015-0016-4, 2015.

[20] Michael Crawford, Taghi M. Khoshgoftaar, Joseph D. Prusa, Aaron N. Richter and Hamzah Al Najada , "Survey of Review Spam Detection Using Machine Learning Techniques", . Journal of Big Data, DOI 10.1186/s40537-015-0029-9, 2015.