# A Survey Paper on Sentiment Analysis of Product Reviews

**Miss. Pawar Bhagyashree D[1] Miss. Shinde Ashwini M[2] Miss. Gaikwad Punam T[3] Miss. Jain Rachana P[4]**

[1,2,3,4]Department of Information Technology

[1,2,3,4]SND Coe & Rc, Yeola, Nashik, Maharashtra, India

*Abstract*— In present days, Internet is the most popular source for getting the information, new ideas, opinions about the available products and different services. People are increasingly attracted towards the various social media like twitter, face-book, amazon etc. As number of people would like to buy the products online, user's reviews or posts are increasing day by day. Also they express their reviews on current affairs, products or services. It is difficult to handle and manage the large amount of data, records and reviews on various products. So, there is a need to classify or perform the sentiment analysis on this large amount of data. The product reviews present in textual format are considered as an input. To work on large amount of reviews data, Support Vector Machine (SVM) is used to classify the review data. We are using Natural Language Processing (NLP) to process the input data. The output of the system is represented as the sentiment expressions such as positive, negative or neutral and it will be shown in the graphical format or in the form of charts.

***Key words:*** Sentiment Analysis, Product Reviews

## I. INTRODUCTION

Basically, the sentiment means the thought, judgment or attitude expressed through the feelings. Sentiment analysis refers to the process of understanding and extracting the sentiments which are expressed in the textual document. The sentiment analysis process includes different steps such as collecting the customer reviews from various social sites or bloggers, processing the collected reviews, classification of review data into different sentiment categories etc.

We know that, Internet is the most popular source for getting the lots of information, getting the new ideas as well as sharing the opinions or views about the different available products and services. As the number of people is sharing their views about different things on the social sites, there is the difficulty to handle and analyze this large amount of data. Hence it is essential to perform sentiment analysis on the data.

## II. LITERATURE SURVEY

*A. Sentiment Analysis of Product Reviews: A Review- By Shiv Prasad [1], Ieee[2017]*

Shiv Prasad represent the sentiment analysis of product reviews which shows the taxonomy of different sentiment analysis techniques. This paper also says that SVM algorithm provides great accuracy than Naive Bayes algorithm. He has explained the sentiment analysis process in five steps as: Collecting the reviews, Preparation of data, Review analysis, Sentiment classification, Display the output in summarized format. But this system is lacked towards fake reviews detection.

*B. Sentiment Analysis using Product Review Data- by Xing Fang [2], Springer [2015]*

In this system, one of the basic problems in sentiment analysis process i.e. the problem of sentiment polarity categorization has been overcame. The general process of sentiment polarity categorization is proposed with detailed explanations. Input data used in this system is the reviews data form Amazon.com. Both sentiment level categorization as well as review level categorization processes are carried out with confirmed results. But the classification of specific star-scaled reviews is difficult in this system. Also the system is limited in sentiment analysis of the reviews which contain implicit sentiments i.e. the sentiments which are represented in neutral words.

*C. Opinion Mining & Sentiment Analysis on Online Customer Review-By Santosh Kumar K. L. [3], IEEE [2016]*

This system focuses on performing the sentiment analysis on product reviews from the social site like Amazon on which customers are able to share their reviews freely. The sentiment analysis is performed on only three reviews from single site using three different algorithms. The three different algorithms used for classification of reviews are - SentiWordNet algorithm, Naive Bayes algorithm and Logistic regression technique. The limitation of this system is that it can work for single site only.

## III. PROPOSED SYSTEM

*A. Architecture:*

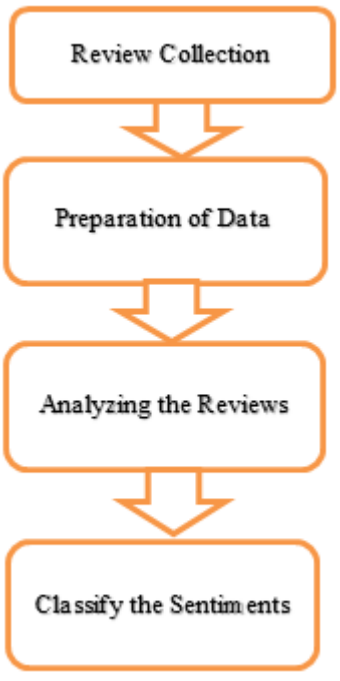


Fig. 1: A Process Model for Sentiment Analysis

 

*B. Sentiment Analysis Process:*

The sentiment on various products is done step by step as follows:

1) Review collection: It is the process of gathering the reviews on various products buy from the different websites.
2) Preparation of Data: The gathered reviews are either in structured or unstructured format. There are multiple reviews on multiple products which leads to big dataset which can be used in sentiment analysis. Sometimes, reviews also contains unnecessary data which gets extracted (tags in HTML, URL information).
3) Analyzing the Reviews: Review analysis means remembering the important characteristics of the reviews then identify the necessary information from the reviews or opinions. The analysis of reviews contains somewhat computational tasks which highlights important data.
4) Classify the Sentiments: Sentiment classification is done in machine learning approach and lexicon based approach. As we know that, machine learning approach is classified into supervised learning and unsupervised learning. Most common or usable classifier are probabilistic and Naïve Bayes classifiers.

## IV. Methods of Sentiment Analysis

*A. Machine Learning Approach.*

– Supervised Learning.
– Unsupervised Learning.

*B. Lexicon Based Approach.*

– Dictionary Based Approach.
– Corpus Based Approach.

## V. Algorithms

1) Support Vector Machine (SVM):- Support Vector Machine is an algorithm in which set of training examples each of which are classified into their corresponding classes is given and the algorithm constructs a model that decides whether a new example comes under one class or the other class.
2) Naive Bays Classifier:- Naive Bayes Classifier is an easy technique to construct the classifiers i.e. models which constitute class labels to different problem instances. This algorithm having the advantage that it only needs a small number of training data to predict the parameters required for classification.

## VI. Applications

*A. Online Commerce:*

The most common use of this system is in e-commerce activities. The social sites allow the customers to share their opinions about shopping and the quality of the products. The system provides overall summary for the products and various features of products by assigning the rating scores. Customers can easily view whole feature information about the product.

*B. Voice of Market (VOM):*

Voice of the Market is about determining what customers are feeling about products of competitors. Accurate and timely information from the Voice of Market helps in gaining competitive advantage and new product development.

*C. Voice of the Customer (VOC):*

Voice of the Customer is concern about what individual customer is saying about products or services. It means analyzing the reviews and feedback of the customers. VOC is a key element of Customer Experience Management. VOC helps in identifying new opportunities for product inventions.

*D. Brand Reputation Management:*

Brand Reputation Management is concern about managing your reputation in market. Opinions from customers or any other parties can damage or enhance your reputation. Brand Reputation Management (BRM) is a product and company focused rather than customer.

## VII. Conclusion

In this paper we are come to know that sentiment analysis plays a vital role for making the decisions about the products. It is also required to focus on certain quality based features of the various products while analyzing the reviews.

There is a wide scope for improving the measures of decision performance in the next versions. Sentiment analysis technique can be enforced on any system which follows the data mining standards. Though the sentiment analysis algorithms are improving rapidly and providing the better quality output, there are many difficulties in identifying the fake reviews. Some people write the fake reviews in such a way that they look like the real reviews and no one can detect their actual motive. Hence the fake review identification is also the significant site that needs the deep data mining technique.

## References

[1] Shiva Prasad T. K., *"Sentiment Analysis of Product Reviews: A Review"*: International Conference on Inventive Communication and Computational Technologies (ICICCT 2017).
[2] Xing fang, *"Sentiment analysis using product review data"*: Fang and Zhan Journal of Big Data (2015).
[3] Santosh Kumar K. L., *"Opinion Mining and Sentiment Analysis on Online Customer Review"*: 2016 IEEE International Conference on Computational Intelligence and Computing Research.
[4] Cambria, *"Affective Computing and Sentiment Analysis"*: 2016 IEEE Published by the IEEE Computer Society.
[5] Preslav Nakov, *"Developing a successful SemEval task in sentiment analysis of Twitter and other social media texts"*: Springer (2016)

 