# OPINION ANALYSIS USING INCREMENTAL MACHINE LEARNING

[1]Sagar Agarwal, [2]Aafra Shaikh, [3]Aishwarya Sharma, [4]Neetu Yadav
[1]Student, [2]Student, [3]Student, [4]Student
[1]Information Technology,
[1]Shree L.R. Tiwari College Of Engineering, Mumbai, India

***Abstract :*** As there is an exponential growth of social network and due to large usage of social media, there is an increasing demand for data in the web for the users which leads to recent trends and ideas in the field of research. The users will be eagerly using these data for the future purpose and get information about the opinions of others thus there is a need of automatic summarization of opinion of the web users. Opinion Mining is a process of extracting and analyzing people's Opinions regarding the given object and Sentiment Analysis explains about the hidden sentiments in the opinions. Sentimental Analysis is reference to the task of Natural Language Processing to determine whether a text contains subjective information and what information it expresses i.e., whether the attitude behind the text is positive, negative or neutral. This project focuses on the aspect based feature extraction of product review using document level sentimental analysis. In a document level classification the main intention is to classify an opinion in the whole document as positive or negative. It speculates entire document as a single unit and it's polarity is decided accordingly

***Index Terms*** **- Opinion mining, sentimental analysis***;*

## I. INTRODUCTION

The breakout of the e-commerce retailers has led to an increase in the reviews given by the customers for the products sold on these websites. Before buying a product a person thinks "What people think?‖. He will have various questions regarding the product like "Which one to buy?‖ etc. The answers for these questions can be found in the reviews which are posted by other buyers on the respective websites but the number of reviews found is overwhelming and difficult to analyze each and every reviews which are expressed in different ways by different customers.
The project Opinion Analysis using machine learning 'helps to provide individual rating to each peripheral and based on these the final rating is given by the system with the help of the reviews on various parameters like cost, performance, etc.

## II. PROJECT SCOPE

The project aims to ease the process of collecting the massive amount of data, so it will be easy for the user to look out for their requirement as per the requirement. In many ecommerce websites wide range of reviews and opinions are available for each product they sell .Through this crucial data information can be extracted to rate the product whether it is good to buy it or not. The project uses various algorithms from the significant domain called opinion mining like part-of-speech tagger (POS), Lexicon algorithm, and Naive Bayesian, etc to extract the information from the data available on the web. Then this information is clustered using clustering algorithm and segregates the data based on its polarity to give appropriate rating to it, when user inputs the requirement .Thus we introduce the imputation method to calculate final rating of a system to help end users to judge about the product.

## III. ABBREVIATIONS AND ACRONYMS

POS- parts-of-speech.

SVM- support vector machine.

## IV. PROPOSED SYSTEM

In today's world of E-commerce, there are too many products to choose from and an average user can get easily confused that what he needs and may also end getting wrong products. The initiative is to ease the time and money of the user is our soul requirement. This will eventually reduce the user stress and will help the user to get the right products easily.
.

**A. WORKING MODEL DIAGRAM**

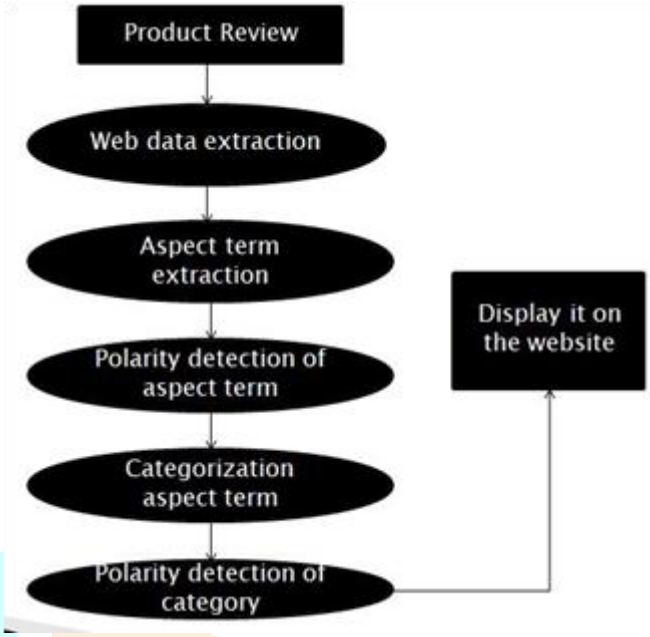


Figure A – Working model Diagram

**B. TECHNICAL OVERVIEW**

The pycharm IDE is used to get results of the sentiments where amazon scrapper, vador sentiments algorithms are used

**Amazon scrapper** – Amazon product scrapper is a great help as it captures all the information of the product/products such as product name, model no., its information, selling details and shipping price. It comes with a One Screen Dash panel which makes possible for you to view all the information on single screen. This dashboard also reveals all the produced keywords, records and elapses. This supply also helps in easy controlling and operation.

**Vader sentiments** – VADER (Valence Aware Dictionary for sentiment Reasoning) is a model used for text sentiment analysis that is sensitive to both polarity (positive/negative) and intensity (strength) of emotion.

**Sentiment Analysis** - Sentiment analysis is a type of data mining that measures the inclination of people's opinions through natural language processing (NLP), computational linguistics and text analysis, which are used to extract and analyze subjective information from the Web - mostly social media and similar sources. The analyzed data quantifies the general public's sentiments or reactions toward certain products, people or ideas and reveal the contextual polarity of the information.

**C. LIMITATION OF EXISTING SYSTEM**

Not all reviews are extracted and the sentiment of reviews with compound nature is difficult to find. Analysis of human sentiments can be accurate up to 70%

**V. FUTURE WORK**

- Integration of specification based rating in current algorithm for more accuracy.
- Smart suggestion based on the previous visit of user.
- Easy to view product reviews.

**VI. RELATED WORK**

Whenever a user buy any product he/she always look for others opinion on that product But on online websites there are on an average 500 to 1000 comments on each product. For a user to read all this comments is not feasible .So the user decides the an opinion on the product just by reading 2 or 3 comments which may not be correct review for the product. This becomes a bigger problem when we shop spare products online where our total analysis about the product is based on others opinion Basically in our system we extract the opinion from the website Then from this data we extract the aspect terms and try to find the sentiments of that aspect terms. These sentiment are useful to find the emotions i.e what user is trying to express. Now these aspect terms are

categorized and then the overall polarity of that category is calculated.These overall polarity is then displayed on the website which helps the user to find which feature of that product is good or bad in single page He/she does not need to read all the comments before buying the product

## VII. REFERENCES

Topic : Web data Extraction techniques:
A review N.V Kamanwar , S.G kale
Year: 2016
https://scihub.io/http://ieeexplore.ieee.org/document/758310 /

Topic : Comparative study of machine learning techniques in sentiment analysis
B.K. Bhavitha , Anisha P. Rodriges ,
Niranjan N. Chiplunkar
Year: 2017
https://scihub.io/http://ieeexplore.ieee.org/document/797519 1/

## VIII. CONCLUSION

This project exploits the association between the opinion expression and then form a coherent review about the aspect that are classified from the sentence using algorithms.Also the sentiments of the aspect term are gained which helps in classifying the product or its aspect terms as good,bad or neutral.