

Research review of User's behavior on E-Commerce based on Opinion Mining Approach

Mr. Manish R. Visa,

Ph. D. Scholar, Department of Computer Science, Gujarat Vidyapith, Gujarat

Prof. Dr. Dhiren B. Patel,

Professor, Department of Computer Science, Gujarat Vidyapith, Gujarat

ABSTRACT:Opinion mining includes different fields like natural language processing, text mining, decision making and linguistics. It is a type of text analysis that classifies the text and makes decision by extracting and analyzing the text. Opinions can be considered as positive and negative and measures the degree of positive or negative related with that event (people, organization, social issues). So, it's basically people's opinion study, study of feelings and evaluations in the course of any social issue, people or entity. After reviewing so many papers there is lot of research needs to be done to identify the fake reviews. This paper focuses on the same.

Keywords: *Opinion mining, Sentiment Analysis, Machine Learning, Natural Language Processing*

INTRODUCTION:

Opinion mining is become an extremely interesting research area due to the obtainability of a huge volume of user-generated content in review sites, forums and blogs. Opinion mining has applications in different areas like in market research, to decision making, to advertising. With the help of opinion mining, companies can estimate the extent of product acceptance and can devise strategies to improve their product. Individuals can also use opinion mining tools to make decisions on their purchasing by comparing competitive products not just based on provisions but also based on user experience and public reviews.

This research focuses on collecting the online user's reviews for any given product or services that user has bought and does the analysis on the reviews based on the sentiments. This includes removal of irrelevant and non-useful reviews and filters the useful reviews. The outcome of the above process would give the benefit to the business/manufacturer in terms of taking strategic decision to enhance the business based on the customer's demand. It also gives the very useful data point to the manufacturer.

Class	Review of Product	Feature of Product
Positive	I bought a Moto G5 S Plus. It has good picture and voice quality. Battery life is also good.	Picture & Voice Quality, Battery Life
Negative	Don't ever buy Micromax mobile as it gets hanged frequently	Processor

Table 1: Review and Feature of Product

E-Commerce websites provides a mechanism where customer can give review for the products that they have bought in positive or in negative way. Negative reviews also need to be considered for taking the decision about enhancement of features/services for particular product.

Customers can take decision based on the following aspects:

- [1] Number of star ratings
- [2] Review's tone in positive or in negative
- [3] Consideration of the product features described in the review
- [4] Important factor which will be used to help customer to buy product based on review
- [5] Review's Authenticity
- [6] Number and age of reviews

Forums, blogs, micro blogs, review sites are the major data sources for opinion mining.

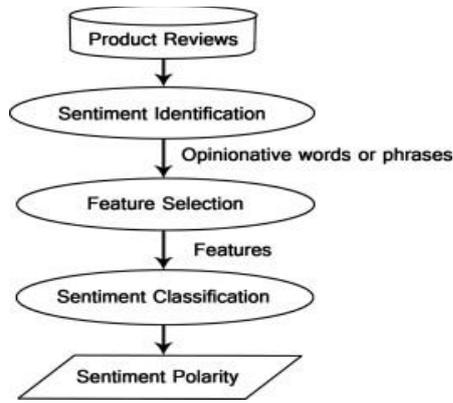


Fig. 1: Diagram of Opinion Mining Process (Source: www.google.com)

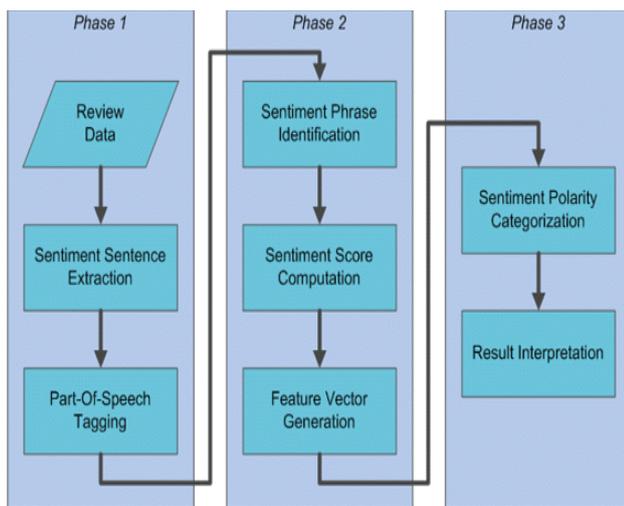


Fig. 2: Sequential steps of Opinion mining process (Source: www.google.com)

LITERATURE VIEW

Literature review has been conducted based on the reviewing lots research paper in the context of the opinion mining.

The author proposes an approach, based on linguistic and statistic analysis of comments written by client on ecommerce website forums, to extend the database for user identification and behavior analysis. The objective is to define the polarity and intensity of opinion written in natural language. [1]

In this paper, author discussed about an attribute level decision support prediction model that has been developed to provide an influential ecommerce platform to the customers. [2]

It uses algorithms to transform free flow text (unstructured) into data that can be analyzed (structured) by applying Statistical, Machine Learning and Natural Language Processing (NLP) techniques. [3]

Aspect based opinion mining approach used First, noun phrases algorithm was used to get all the aspect term of a review sentence. Secondly, the sentiment algorithm was applied on the result of the noun-phrase algorithm and also applied on adjectives and on adverbs. Finally, using relative importance algorithm important aspects were presented to the user. [4]

The proposed method uses Apriori algorithm for feature extraction based on the customer reviews. The classification is done on product features based on unsupervised SentiWordNet method. [5]

It automatically extracts the reviews from the website. It also uses algorithm such as Naïve Bayes classifier, Logistic Regression and SentiWordNet algorithm to classify the review as positive and negative review. [6]

Opinions on specific topic are inevitably dependent on many social effects such as user preference on topics, peer influence, user profile information. It uses Words of mouth (WOM). [7]

It proposes a novel framework for recommendation system by utilizing an FPIntersect algorithm. The system uses the information obtained after analyzing the opinions of each user from the user comments. [8]

This paper presents a novel application of existing automatic extraction, sentiment analysis and summarization techniques for product review annotation classification (i.e., to discover product as recommended/not recommended in a review), product attribute popularity tracking, extracting aggregate positive vs. negative opinion. [9]

A novel technique that uses a multi-dimensional trust evaluation model, for computing comprehensive trust scores for sellers in e-commerce applications. The system computes dimension trust scores and the dimension weights. [10]

TECHNIQUES USED IN OPINION MINING

Based on the current trends, one or more methodology can be adopted from following methodologies.

Methods:

It indicates algorithms to transform free flow text (unstructured) into data that can be analyzed (structured) by applying Statistical, Machine Learning and Natural Language Processing (NLP) techniques

Natural language processing:

To automatically read reviews and used Naive Bayes classification to determine the polarity of reviews.

There are different approaches of classifying the review sentences using supervised method and unsupervised method of Machine Learning Technique.

Supervised methods describe the Classification and Regression concept.

- Classification: It consist categorical variable like disease or no disease, red or blue.

- Regression: It consist real value like percentage, height, width, age etc.

- Following is the list of supervised methods algorithm

- 1) Naive Bayes Classification

- 2) Support Vector Machine

- 3) Maximum Entropy Rule

Unsupervised methods describe the Clustering and Association Rule concept.

- Clustering: It describes the group like based on the customer's purchasing behavior

- Association Rule: It is used to identify the pattern or discover the rules from your dataset.

- Following is the list of supervised methods algorithm

- 1) K-means

- 2) Apriori Algorithm

Naive Bayesian and SVM method requires trained data set of opinion words. SentiWordNet is an unsupervised method used for classification of review sentences. This work uses unsupervised method using SentiWordNet and takes AAVN (Adjective, Adverb, Verb, Noun) from the review sentence as opinion words. Usually adjectives, adverbs and verbs are used to express the opinion words.

Stop word removal algorithm:

Stop words are word that has small meaning apart from the other words. The most common stop words are "the," "a," "an," "that," "those" and so on. These words need to be removed.

POS Tagging:

POS tagger is part-of-speech tagging, which is used to tag each and every word used in the review sentence to its proper part of speech i.e. it tags words like as noun, adjective, verb, adverb, conjunction, preposition and interjection. POS tagging of words is essential to find features and opinions expressed by the customer about the product in the review sentences.

RESEARCH CHALLENGES AND GAPS IN OPINING MINING

How to identify the Fake Reviews?

- The customer who bought the product can only be give the review either through the OTP (which will be send to his mobile number) or through the link which has been sent to his email. After clicking on the link, he/she would be able to give the review.

- There should be mechanism to validate the review against rating like customer can't give outstanding feedback in the case of he has given 1 rating and bad feedback in the case of 5 rating.

- This would be the mechanism to avoid the fake reviews.

How to identify abbreviations used in reviews?

- There should be database/files which will holds most commonly used abbreviations and based on that reviews can be determine in the terms of positive, negative or neutral.

Orientation of Opinion Words: [1]

- It could be difficult to interpret word based on situation. For e.g. The size of the camera is small. Here small is used in positive sense but if customer says like the battery life of mobile is small then it is used in negative sense.

Group of Synonym Words: [1]

- It is one of the challenging task to relate same feature but in different style like quality of mobile excellent and quality of mobile is very good gives the same feature.

Domain dependent nature of sentiment words: [1]

- One feature may have very good performance in one domain but at the same time it will have bad feature in another domain.

Consideration of Social media to extract the product reviews: [1]

- Not only from the e-commerce websites but there should be mechanism to retrieve the product review from social media.

CONCLUSION AND FUTURE ATTEMPTS

Opinion mining is used to determine and extract the subjective information from the text and it is also called the sentimental analysis.

Sentiment analysis

Sentiment analysis systematically identifies, extracts, quantifies and studies affective states and subjective information making use of natural language processing, text analysis and computational semantics. Sentiment analysis is mostly used to identify the customer's buying intension for particular product or service based on the positivity and it would be useful for marketers as well. Product's positivity or negativity can be analyzed by Admin. In star rating

scheme, it is easy to count the stars. In text review scheme, it can be extracted and matched with the database.

I have gone through the many papers in order to perform review on opinion mining approach.

Types of fake reviews:

Polarized reviews: With high rating of the product and polar opposite review in the review text

Paid reviews: Writers paid to write high quality good or bad reviews about a product. Hard to distinguish.

How to tackle polarized reviews:

Indicators: Positive sentiment in rating and negative sentiment in review (and vice versa)

Initial idea is to build a classifier (get input data from document and put the document in one of the N classes) that can classify the sentiment in the review text as positive, neutral or negative. One can achieve this with Multinomial Naive Bayes, SVM or even a neural network.

From the output of our classifier, we can tally it with the review score and identify polarized ones. We expect the positive reviewer to give good rating (e.g. above 2.5) and negative reviewer to give bad rating. (e.g. below 2.5)

The mean and standard deviation of reviewers rating can also be used for determining if a reviewer is more inclined to give positive review or vice versa.

REFERENCES

- [1] Surya Prakash Sharma, Dr Rajdev Tiwari, Dr Rajesh Prasad. Opinion Mining and Sentiment Analysis on Customer Review Documents - A Survey. International Journal of Advanced Research in Computer and Communication Engineering, 2017.
- [2] Sujoy Baga, Manoj Kumar Tiwaria, Felix T.S. Chanb. Predicting the consumer's purchase intention of durable goods: An attribute-level analysis. Journal of Business Research, November 2017.
- [3] Chandrasekhar Rangu, Shuvojit Chatterjee, Srinivasa Rao Valluru. Text Mining Approach for Product Quality Enhancement (Improving Product Quality through Machine Learning). IEEE 7th International Advance Computing Conference, 2017.
- [4] K. Srividya, K. Mariyababub, A. Mary Sowjanya. Mining Interesting Aspects of a Product using Aspect-based Opinion Mining from Product Reviews. International Journal of Engineering, November 2017.
- [5] Anisha P Rodrigues, Dr. Niranjan N Chiplunkar. Mining Online Product Reviews and Extracting Product features using Unsupervised method. IEEE 2016.
- [6] Santhosh Kumar K L, Jayanti Desai, Jharna Majumdar. Opinion Mining and Sentiment Analysis on Online Customer Review. IEEE 2016.
- [7] Nitu Kumari, Dr. Shailendra Narayan Singh. Sentiment Analysis on E-commerce Application by using Opinion Mining. IEEE 2016.
- [8] P Devika, R C Jisha and G P Sajeev. A Novel Approach for Book Recommendation Systems. IEEE 2016.
- [9] Gaurav Dubey, Ajay Rana, Naveen Kumar Shukla. User Reviews Data Analysis using Opinion Mining on Web. 2015 1st International Conference on Futuristic trend in Computational Analysis and Knowledge Management. 2015.
- [10] Emayakumaari T, Ananthi G. Mining E-Commerce Feedback Comments for Trust Evaluation. 2015 IEEE International Conference on Engineering and Technology (ICETECH), March 2015.
- [11] Quinliang Miao, Quidan Li. An Opinion Search System for Consumer Products. IEEE 2015.
- [12] Ms. Ashwini Rao, Dr. Ketan Shah. Model for Improving Relevant Feature Extraction for Opinion Summarization. IEEE 2015.
- [13] Venkata Rajeev P, Smrithi Rekha V. Recommending Products to Customers using Opinion Mining of Online Product Reviews and Features. International Conference on Circuit, Power and Computing Technologies [ICCPCT] 2015.
- [14] Dongzhi Wang, Xinwei Yan, Huimin Wang and Xiu Li. A Conceptual Framework of E-Commerce Supervision System Based on Opinion Mining. International Conference on Services Science. 2015.
- [15] Hsien-You Hsieh, Shih-Hung Wu. Ranking Online Customer Reviews with the SVR Model. IEEE 16th International Conference on Information Reuse and Integration. 2015.
- [16] Katowice, Poland, Katarzyna Wegrzyn-Wolska, Lamine Bougueroua. An Opinion Mining Approach for Web User Identification and Clients' Behaviour Analysis. ResearchGate. 2014.
- [17] Deepali Virmani, Vikrant Malhotra, Ridhi Tyagi. Sentiment Analysis Using Collaborated Opinion Mining. ResearchGate. 2014.
- [18] T. K. Das, D. P. Acharjya and M. R. Patra. Opinion Mining about a Product by Analyzing Public Tweets in Twitter. (ICCCI) International Conference Computer Communication and Informatics. 2014.
- [19] Mrs. Vrushali Yogesh Karkare, Dr. Sunil R. Gupta. Product Evaluation using Mining and Rating Opinions of Product Features. International Conference on Electronic Systems, Signal Processing and Computing Technologies. 2014.
- [20] Prashast Kumar, Arjit Sachdeva, Dhruv Mahajan, Nishtha Pande, Amit Sharma. An approach towards feature specific opinion mining and sentimental analysis across e-commerce websites. IEEE 2014.