A Study on Review Systems in Ecommerce

Sheetal Kandhare[1], Chandrashekhar Pawar[2], Sudeshna Gunjote[3]

[1,2,3]Student, Department of Computer Engineering, Dr.D.Y.Patil School of Engineering Academy, Pune, India

---

Abstract - *The propose of this case study is to explore various online E-commerce websites and online portals to analysis review system and ranking systems on the basis of various technologies and methodology. The key objective of this paper is to find the efficient and reliable way both to analysis reviews and customer preferred data analysis for relevant ranking of reviews. This research work first employee's orderly literature review of methodology. It consists of diversifying methodologies and techniques and analysis.*

***Key Words*: E-commerce, Review System, Data mining,**

Ranking

1. INTRODUCTION

The modern era has a humongous impact on the Ecommerce websites and an important part that affects to all the online shoppers is the Review. E-commerce has a immense impact in today's world all the shopping from small to whole vend is available and people prefer to shop online without going out. In covid-19 pandemic situation the growth of local as well as ecommerce has gain a huge rise in orders. The people would rather go for ecommerce websites such as amazon, flipkart, eBay etc. rather than searching for the same item in different areas and stores. Since gradual increase in demand of ecommerce website the customer need assurance for the product they buy. Questions such as Are the product good? ,Is the battery life good for any particular laptop or mobile?, Does the product has any problems? For such kind of questions review and ratings are always preferred. Various positive and negative review or helpfulness of the review should be correct as per the authenticated customer or buyers. According to oberlo(Oberlo,2021), E-commerce websites produces an estimated that there will be 2.14 billion global digital buyers in 2021and more than 60% of the customer prefer to buy an item based on the reviews and rating of various users. BrightLocal.com also states that 90% of consumers read ten or less reviews before making their purchase decisions.(Sunil Saumya et.al,2018)

Web mining(Wenpu Xing,2004) and data analysis is used to discover past behavior of users and future preferences to view pages. The review recommendation and ranking of reviews based on various machine learning approaches.

Machine learning techniques for analyzing reviews based on their lexical semantics features in order to differentiate or categorized it in balanced review or unbalanced review.(Helpful review or Not an helpful review)(Ralf Krestel et.al,2015).Review recommender systems that uses machine learning algorithms also has the abilities to classify the review as helpful or not, or to identify the helpfulness of the review. Some popular commercial Ecommerce websites such as amazon and eBay uses this type of approach and even has ranking based on the review given by a particular verified user by their website. For eg. Amazon states whether the reviewer is a verified user and how helpful was his review to other relevant customers. There are approaches that are either manual or automated for predicting the helpfulness of the given review. It is very essential to identify whether the given review is easily understandable, is the review negative review or a positive, what accepts the reviewer has focused on. The Very first an important step for predicting an review is good or bad is try to gain information form it. Amazon has Question and answer section to make customer engage and to have customers attention and solve their query's.

The online Ecommerce is the bridge between the business and the customer. Today most of the Customer connect to the websites and product by seeking confidence through reviews and ranking of the product. Customer Review plays a important part for the ecommerce. For a single product hundreds of reviews are posted on the site. Identifying, reorganizations the helpfulness of the reviews and sort or rank the reviews accordingly. Many of aspects such as time posted or ranking of the reviewer and votes for certain reviews.

People prefer to read helpfulness reviews and ranking of the reviews accordingly. All the filtration methods and sorting algorithms to filter the top reviews referred for the customer to gain confidence for particular purchase decisions.

2. REVIEW RECOMMENDER SYSTEM.

The customers on the online e-commerce as well as in offline purchase the products and items as per the reviews of past experiences. The major impact of reviews is for all the ecommerce websites the customer may end up buying the wrong products so, to solve this machine learning algorithm techniques are featured to analyze the helpfulness or do gain the information from the reviews and analyze the best reviews that can help other users to seek the confidence. The probability of all reviews being helpful is very low. Due to the fact we can prioritize it accordingly. Therefore, researchers are improving the techniques.

A. Content-based filtering technique.

Content-based filtering is Item similarity content based. It takes Content are taken under consideration for filtration

of reviews. Document stores are taken then filtered according user interest prolific using utility function. This based binary classifiers. Filtered reviews are then accepted by the function.

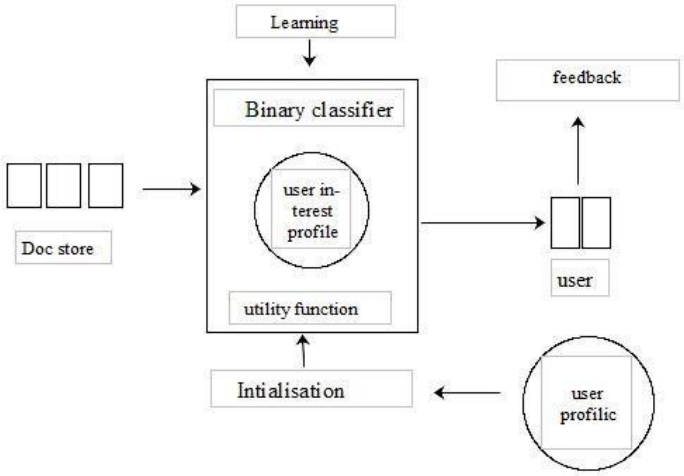

Fig. Content-based recommendation system

The content-based recommendation system is based on a single user's similarity of items, it accumulates all the datasets and user's profile interests and analyzes a particular threshold for acceptance of the data. The learning model is for learning the information and users' behavior, Does the user buy the particular product? Or does the user went through recommended products?

This will indicate the product or the recommended thing is relevant or not for that particular user evaluation takes place by the utility functions.

Linear utility = 3*#Good-2*#Bad

B. Collaborative filtering technique

Similar to the content-based filtering technique the only replacement here is the item similarity by the user's similarity.

The items decided by the users are taken into the course and the person's interest is analyzed and recommended. Thus, using this for reviews will be according to the basis on all the product specifications such as batteries of the phone (mobile), cameras of the mobile, frequent views of the particular reviews, top reviews measures. Users get broader exposure and this increases the possibility of continuous usage or purchase of the product.

C. Memory-based filtering techniques

The product /items that are rated before by the customers are helpful for the recommendations of neighbors of the user's through item and user-based filtering techniques. The users past preferences and reviews would help for further neighbor's recommendation of products or items. (Bansari Patel et.al,2018)

1. User-based filtering

User-based filtering accepts reviews or data that are similar to the users. No items are taken under consideration for filtering, comparing the reviews, and ranking between users takes place.

2. Item-based filtering.

Item-based filtering takes the user's items under consideration, then after learning and observing the datasets, relevant items are recommended to the given users. (Bansari Patel et.al,2018)

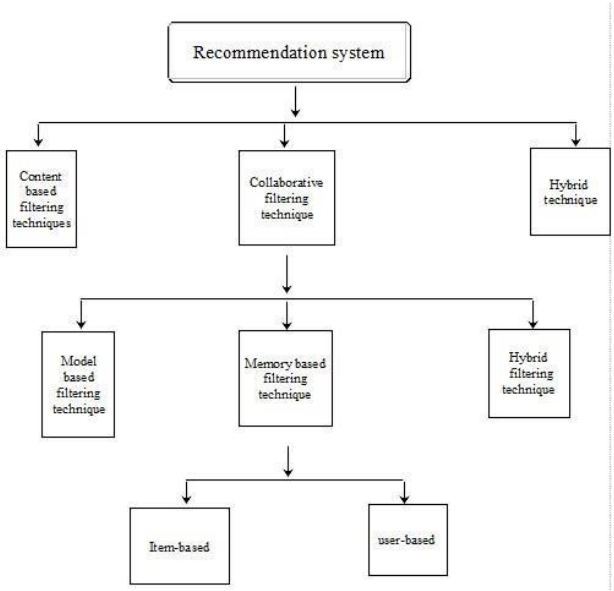

Chart -1: Types diagram

(Bansari Patel,2018) **D.**

Model-based filtering.

Model-based filtering techniques are based on the user's model of rating the reviews and rating. A model of user's review and rating is created. Algorithms in this technique take probabilistic approaches and compute the expected value of prediction i.e. helpfulness of the review.

E. Hybrid filtering techniques

The hybrid filtering technique is a combination of different filtering techniques. Two or more filtering techniques can be combined to make several functions and then to make relevant dataset available.

3. WEIGHTED PAGE_RANK ALGORITHM

The Weighted PageRank algorithm is a widely used algorithm in the field of ranking items. It is normally observed that more linkages than other web pages would

have it. It estimates how important the review is or not, The information gained by a particular review is helpful or relevant for the item. Helpfulness of the review is calculated either negative (Helpfulness) or positive (Helpfulness).

The Page Rank algorithm (Wenpu Xing,2004), states the importance of the link to it and then another page. So, the PageRank algorithm takes backlinks into consideration. If 'X' visits 'A' then visit next to 'B', again 'X' went to page 'A' so this observation is taken and no. of backlink to the A are more. Therefore,' A' has the highest-ranking of the page.

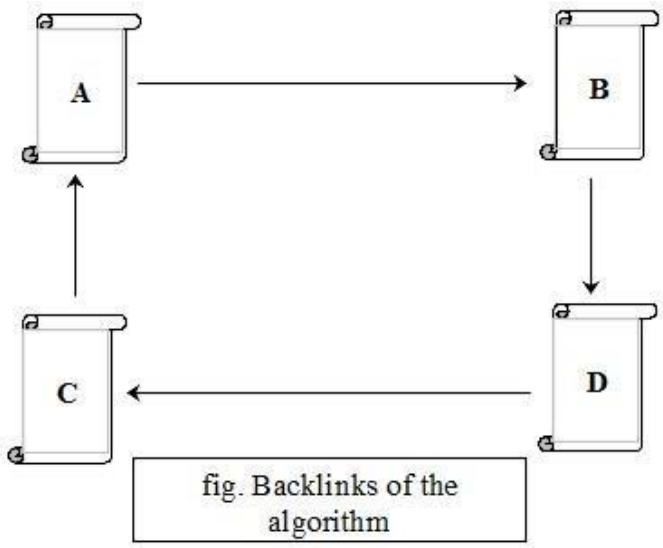

fig. Backlinks of the algorithm

Higher the backlinks counts, higher the importance, evaluation of reviews using WPR algorithm is implemented as:

1. Finding the helpfulness of the review.

2. Observing or Learning the counts on that

review.(Helpfulness of the review for other peoples)

3. More the helpfulness, higher the priority.

4. Ranking the review, according to the helpfulness that may be positive or negative helpfulness.

Reviews can be classified as:

- Very significant reviews: These are the reviews which have all the positive /negative description.
- Significant review: This review has less information gain about positive /negative descriptions than very significant reviews.
- Weak reviews: These reviews do not have any helpfulness value and are not properly described.
- Irrelevant reviews: These reviews are those which do not have any relevance to the item or product purchased.

4. Query- Dependent Ranking Approach

A ranking model that is based on the approach "Learning to rank" (Lian-Wang Lee et.al, 2009) which attempts to score all the documents in a single function.

Recently several "Learning to rank" algorithm has been developed and formulated as a binary classification that can be divided into two parts as not helpful or helpful(Positive or negative review rank).

This is a simple similarity measure approach to evaluate the similarity between two reviews.

When a testing query is performed to the given query, a similarity measure is applied. Then, the queries are put together to obtain the overall score and retrieve the weight gain. At last rank the review accordingly.

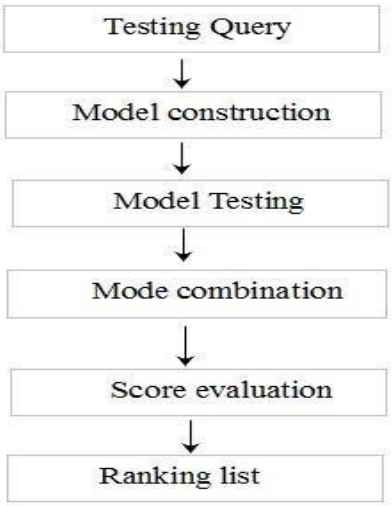

fig.Query Approach(Lian-Wang Lee et.al,2009)

5. Text-Summarisation Approach

Item reviews pose critical data that makes the consumer confidence seeker and all the review system approaches are mainly on the binary form or categorically classified, It has been observed that textual data is been ignored. (J. Zhan et al, 2007)

A typical preference based on marketing an item is based on review according to the filtering of the consumer.

Most of the existing works are performed with opinion mining, data mining, and different approaches such as novel graph-based summarization. [Ganesan et al (2010) (Wang et. al (2013)] developed a web-based review summarization, then concept level approach [(Lirot et. al(2015))].

Automatic text summarization has explosive growth in last years. Example Amazon.com, walmart.com. Majorly two types of techniques are used as statistical approach and linguistic approach.

The clustering summarization has increased its domain, this separates the documents into different groups and then summarizes each of them. Many clustering algorithms are implemented by (Michael et.al Gamon,USA) (J. Zhan et al,2007)

Tropical structure (Zhan et al, 2009) approach that proposed a structure consisting of a list of different reviews based on various significant topics that are extracted from the document set. Then summarized each topic's reviews on different topics.

[Abdi et al(2018)] developed a summarize the user opinions and reviews based on machine learning-based approaches. This calculates the sentiment score and determines the salient sentence on statistical and linguistic knowledge which provided significant performance.

6. ARTIFICAL INTELLIGENCE FOR REVIEWS

Artificial intelligence or knowledge base techniques are more advanced which also have the capability to predict and as the classical approach to gain information.

Neural network which is effective predictive modeling techniques, the neural network is unsupervised learning method and is used in complex data software engineering perspective in intelligence systems.

A decision tree (Sang Jun Lee,2001) approach is a classification approach that is based on the greedy approach, CART (Sang Jun Lee,2001) is the technique used for the classification of data set.

Predictive analysis ecommerce by [Predictive analysis, (Jagatjyoti, Ashish)2018] a predictive model for ecommerce sites that uses Native language toolkit, stemming and filtering approach for ranking and rating generation.

7. DATA MINING APPORACH

The growth is data sets have been a rapid explosive growth of database results in developing new technologies and methods to make expository easy.

Data mining is an intelligent and beneficial approach to overall extract or gains information through a database. Therefore, with the increase in E-commerce use product reviews increase day by day, Data mining can be used on these reviews to in identify the pattern, helpfulness, and attractiveness of the review. So, that can be ranked appropriately. Data mining and text summarization also implemented by (Deng Bin et.al,2010). A. Based on database:

This approach is based on databases such as relational, object-oriented databases, A DM can extract knowledge from the database and calculate the helpfulness of the review. (Sang Jun Lee, 2001) B. Based on the knowledge:

DM systems have rules that are used to identify or extract information, rules such as deviation analysis, association rules, and clustering.

C. Based on techniques:

Different techniques of DM are categorized according to the data available. Data-driven mining, query-driven mining approach based on mathematical and statistical mining. (Sang Jun Lee, 2001)

8. DISCUSSION

The major highlight of the review system is that the review system must be identified recognized with the help of different methodologies, the more the helpfulness of the review is calculated, the more is the review prioritized and ranked according.

The primordial techniques were less efficient than the latest methodologies we have now, It is quite substantiated that review has a major impact on the customers and a bad review can lead the customer to fallacious decisions.

The major two approaches stated in section 6 and section 7 have to lead the booming technology in use. Data mining has various techniques which lead the reliable ways to gain information and rank accordingly, with the help of artificial intelligence, it is possible to predict the reviews that are good and bad classification based on different topics of the review can be achieved with it.

Divergent and intelligent approaches are being developed the future works can be implemented on big data and analyzing and predicting and ranking than that manual approaches are still used in many websites, websites such as TripAdvisor, Amazon, Walmart, Flipkart are ahead for the use of technology and preferences and customer satisfaction.

9. CONCULSIONS

We have observed that having and extracting the right information is intrinsic o make the right decisions, Ecommerce has exorbitant reviews, and extricating information from such big data is pretty possible in today's competition in organizations, Although progress in review systems.

This paper presents a review of literature and techniques that can be developed and implemented that can be

developed and implemented for an e-commerce review system. Finally, we conclude the methodologies and the core of each section, and applications for future work.

10. REFERENCES

[1] Data mining techniques and applications – A decade review from 2000 to 2011 Shu-Hsien Liao ⇑, Pei-Hui Chu, Pei-Yuan HsiaoM. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[2] A review of data mining techniques – Sang Jun Lee, Keng Siau, University of Nevbraska-Lincoln,Nebraska, USA

[3] E-Commerce Reviews Management System Based on Online Customer Reviews Mining Deng Bin School of mgmt. and Econ. Univ. of Elec. Sci. & Tech. of China Chengdu, P.R.China ,Shao Peiji ,Zhao Dan [4] Gather customer concerns from online product reviews – A text summarization approach Jiaming Zhan a , Han Tong Loh a , Ying Liu b,*

[5] Pulse: Mining Customer Opinions from Free Text Michael Gamon, Anthony Aue, Simon Corston-Oliver, and Eric Ringger

[6] Predictive Analysis of E-Commerce Products Jagatjyoti G. Tuladhar, Ashish Gupta, Sachit Shrestha, Ujjen Man Bania and K. Bhargavi

[7] Methods of Recommender System: A Review Bansari Patel Department of Computer Engineering CGPIT, Uka Tarsadia University, Bardoli, India. bansari2404@gmail.com Palak Desai Department of Computer Engineering CGPIT, Uka Tarsadia University, Bardoli, India. palakv.desai@utu.ac.in Urvi Panchal

[8] Ranking online consumer reviews Sunil Saumya, Jyoti Prakash Singh, Abdullah Mohammed Baabdullah, Nripendra P. Rana, Yogesh K. Dwivedi

[9] A Query-Dependent Ranking Approach for Search Engines Lian-Wang Lee1 Jung-Yi Jiang1 ChunDer Wu2 Shie-Jue Lee [10] Weighted PageRank Algorithm Wenpu Xing and Ali Ghorbani

[11] Diversifying customer review rankings Ralf Krestel a,∗ , Nima Dokoohaki b a Department of Computer Science, University of California, Irvine, USA b Decisions, Networks and Analytics (DNA), Swedish Institute of Computer Science

[12] Review of current data mining techniques used in the software effort estimation – Julius olufemi Ogunleye

[13] A Review Of Usability and security Evaluation Model of Ecommerce -Nur Azimahbt,MohdZarul,FirtriZaaba

[14] https://www.oberlo.com/blog/ecommerce-statistics

[15] Vinod Bharat et al. "A review paper on data mining techniques", International Journal of Engineering Science and Computing (IJESC), 2016, Volume 6 Issue 5, pp 6268-6271.