

# Investigation of Ambiguity based Sentiment Analysis for Product Recommendation on E-Commerce Portal

Ruchira Muchhal<sup>1</sup>, Mohit Jain<sup>2</sup>

<sup>1</sup> M. Tech Scholar, CSE Department BM College of Technology, Indore, India

<sup>2</sup> Assistant Professor, CSE Department BM College of Technology, Indore, India

**Abstract:** *Rapid growth of internet and their relevant applications make it a backbone of routine life. Hectic schedules and busy routine life make it impractical to visit and purchase product personally. E-commerce based shopping portals are trying to fulfil this gap and helping users to save their time and money. Recommendation system are playing crucial role in this dimension because they help users for easy selection and reduce shopping time. Different studies are under going in dimension of product recommendation for e-commerce portals and improve recommendation process with minimum shopping time and find that most of the recommendation systems only work on product selection frequency and user rating. User's previous purchasing record and other user's sentiments about product can play significant role to make trust on particular product during online shopping. A survey has been done in this dimension on different relevant work to explore the gap area in conventional system and possibilities of solution to overcome the same. This paper provides a survey report on need of opinion mining with respect to e-commerce portals and shopping websites.*

**Keywords:** *Opinion Mining, Sentiment Analysis, Product Recommendation, E-Commerce Portals*

## I. INTRODUCTION

The market of online portals and e-commerce websites for product recommendation is growing with rapid rate. In recent years, many e-commerce sites have been developed like flip kart, Amazon, Myntra etc for online shopping and have become big elephant of this market. People are spending a lot of time on such sites for shopping and planning for purchasing on upcoming festivals. Online portals have become a great opportunity for customer and traders both. Besides all this bright sites they are also lacking with the issue of customer trust on product and exploring similar product according to their expectations. Here, user's rating and reviews can play a vital role and help customer to take decision about purchasing. The problem product rating and review analysis address is that, rating is always available in numeric format so it can be easily quantified but review cannot be quantified. Text reviews are always used to explore user's sentiment and opinion about the product. Sometime user may use product review to give feedback about product quality and service support. Subsequently, users rating and review can also help to decide product rating as well as sellers rating. This work explores that user rating plays a big role in only purchasing. Customer always prefers a good rating product for online shopping. Study of existing systems explore that most of the portals either consider rating to analyze users view point or user feedback to analyze service and support values. Users review can play big role to discover about overall opinion of different users on particular products and help in smooth purchasing. Although, selling and purchasing is natural process but user opinion on different products can also help to analyze about organization's process. Big quantity of negative review on different product may help to discover existing problems and gap area in e-commerce portals and help to improve their performance. Recommendation system is an information filtering system which on the basis of user's behaviour and past transactions presents the list of recommendation. This system of recommendation is very common system which is applied in several applications. On the basis of opinion mining and content based filtering system, data mining can help to explore about product popularity and viewpoint of products. User interest and behaviour can be classified based on different shopping patterns. Study of different existing solution and current applications explore the need of more contribution on review analysis and improve product recommendation policy for e-commerce portals. This paper consists study of existing solution, a brief comparative analysis, gap area in existing solution along with tentative plan to overcome existing problems and improve product recommendation and opinion based review analysis for e-commerce portals.

## II. RELATED WORK

Ashraf Elnagar [1] presented an investigation research work for Arabic language. They used sentiment analysis approach for review analysis and explored that it can help to quantify text review on numeric values. Author considered Arabic language as input language platform and used bag of word methodology to translate them. They used Arabic book's reviews (LABR) dataset for analysis purpose which is available for free. They also performed experimental analysis and observed performance of this sentiment extraction accuracy based on precision, recall and f-score parameters. Mir Riyanul Islam [2] address that how sentiment analysis can help to extract users opinion about product. Author also explored that quantification of text review can help to fix product rating and improve product recommendation approach for e-commerce portals. They address that ambiguity can create big hurdle in line of sentiment analysis so there is a strong need to identify sentiwords from review analysis. A.K. Singh et al. In[3] proposed the occurrence of frequent itemsets by the association rule system, by calculating the frequency of appearance of items. Attractive counts are very essential and important and have a vital role in database. It is used to reduce the size of database. The unwanted and unused data is removed from the database. From the set of heavy rules the unimportant rules are removed to minimize the database size and with this it produces the new set of rule, this is called the attractive counts. S. Rao et al. In[4] described about the new approach for association rule which is based on the amount of used memory, rules of interest, time and number of scan database. To overcome the drawback of traditional apriori algorithm, data mining association algorithm is discovered. For avoiding the iteration of scanning database again and again association rule mining algorithm is used and is very effective for pattern mining. O. R. Zaane et al. In[5] explained about web mining use and its working approach for recommendation of activities for online learning. They address that mining of product information can help for product lead generation along with recommendation. This paper do some needful in the same direction where they used simple mining approach for product recommendation.

R. Srikant et al. In[6] proposed about the above arrived issue and to solve it author described about the combinational algorithm which is implemented as Hybrid algorithm and called as Apriori Hybrid algorithm, this algorithm is completely different from the above described algorithm. Apriori Hybrid algorithm is based on the transaction size and number of products available in database, this algorithm has some exception feature property. Traditional algorithm only deal with the equal priority of all the transactional items present in database.

P. Nagarnaik et al. In[7] described about the exponential growth in data in the recent years. For the advertisement and best recommendation they depend upon the online shopping sites. It is beneficial for both customer and consumer. For customer it is advantageous as they can get better recommendation on the basis of their choice and for the consumer it is advantageous as they are preferred more for doing a good job in recommendation field. Customer's opinion and feedback are more important for the growth of any organization. It is also a difficult task in mining, to extract heavy information from a large database. That is why; a better algorithm is required to mine user interest. Many approaches can be used to generate recommendation, this approaches are content-based filtering, collaborative filtering, hybrid approaches etc.

P Devika et al. In[8] proposed a recommendation system for e-commerce applications. Their study concludes that product recommendation is not possible without having an idea of user's interests, other user's preferences, and their ratings. To provide a better recommendation system, it is necessary to generate associations among products. Since e-commerce and social networking sites generates massive data, traditional data mining approaches perform poorly. Also, the pattern mining algorithm such as the traditional Apriori suffers from high latency in scanning the large database for generating association. They proposed recommendation algorithms using FPIIntersect algorithm and compare book recommendation results with Apriori algorithms. Unique factor of their proposal is involvement of sentiment analysis approach to elaborate relation between user and book. Proposed solution is elaborated in the figure below.

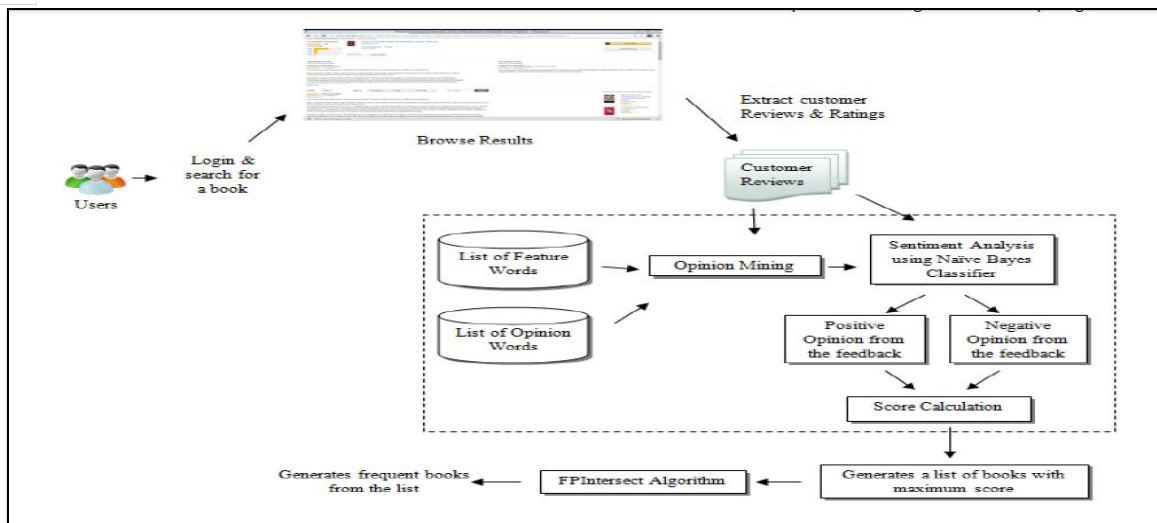


Figure 1: Existing solution by Devika[8]

K. Hong et al. In[9] addressed about the deficiency of previous work, where user information is sensitive and can not be defend to public. Author ‘s research concluded the high accuracy and high satisfaction of rapid use of users profile. Also gives result for the consumption of more time, where user spend lot of time in searching research paper for the specific subject.

J. Yang et al. In[10] proposed an improved algorithm by using feature Apriori algorithm, where a feature of transaction item is added using mining rule and association rule by making it more efficient and reliable

### III. PROBLEM STATEMENT

Study of relevant existing system and relevant research work concludes that product rating and review plays very important role in e-commerce portals. User trust is the biggest factor behind every shopping. Sentiment analysis is much diversified technology that can help to extract user opinion and discover reviewers view point for respective product. Study of existing solution finds that most of the solutions are developed based on rating factor as well as some numeric feedback. Elnagar [1] performed sentiment analysis but on Arabic language. Subsequently, they do not consider any ambiguity analysis or updations of frequent data analysis and integration.

This research work analyzes different research paper to observe about need of text review analysis and importance of sentiment analysis on opinion extraction. Study of a book recommendation research work done by Devika[8] et. al. explore that product recommendation cannot be done without having prior information of user interest and users text review. Subsequently, user’s text review can also help to extract popularity and frequent use of product on online shopping sites. Devika et. al proposed a solution to improve book recommendation policy using FP-intersect algorithm over Apriori Algorithm. FP-intersect algorithm helps to overcome drawbacks of Apriori algorithm for most transactional dataset. FP-Intersect algorithm works on multiple batches and only performs revised analysis on updated information. But it does not perform ambiguity analysis during sentiment analysis. Ambiguity factor can affect a lot on review analysis and may deflect the opinion during finding the user interest. Subsequently, they do not consider rating to quantify opinion of user interest. In the same line they did this task for small data, so expansion of opinion analysis for large data can be a big contribution. The complete study observe that there is big need to implement a solution having capability to perform sentiment analysis along with ambiguity analysis for user text review extraction along with involvement of numeric rating for opinion mining. The complete study conclude that large data processing should be performed along with improved sentiment analysis where text review should not only be observed based on polarity but also on ambiguity factor.

### IV. PROPOSED SOLUTION

Proposed recommendation system will attempt to identify products on the basis of user interest depending upon the popularity factor and their preferences. The similarity might be based on the similarity of items which most of the user prefer or likes or in past made transactions for those items or based on rating and review of that item. The complete work concludes that proposed method will be implemented as Product recommendation system to get better recommendation on the basis of interest. Following steps will be proposed to experimentally analyze the Product recommendation system.

- A. Amazon dataset will be used as input data source for review and rating analysis.
- B. A complete ambiguity based sentiment analysis will be done
- C. Initial module will implement pre-processing of individual user review where each review will be tokenized and lemmatized by Stanford lemmatize and keyword extraction library.
- D. Sentiwordnet list will be used as wordnet dictionary and identification of polarity of individual words.
- E. Polarity analysis of individual words and respective review will be done to convert text sentiments into numeric sentiments representation.
- F. Sentiment weight calculation and sentiment weight for individual review is calculated with applied weight and then senti score is calculated based on user review and rating.
- G. In the next step Frequency Pattern Intersect Algorithm (FP Intersect) is applied on the calculated score to get better recommendation.
- H. Finally, product recommendation will be displayed based on sentiment analysis.
- I. The complete solution has been demonstrated in block diagram which is shown in figure 2;

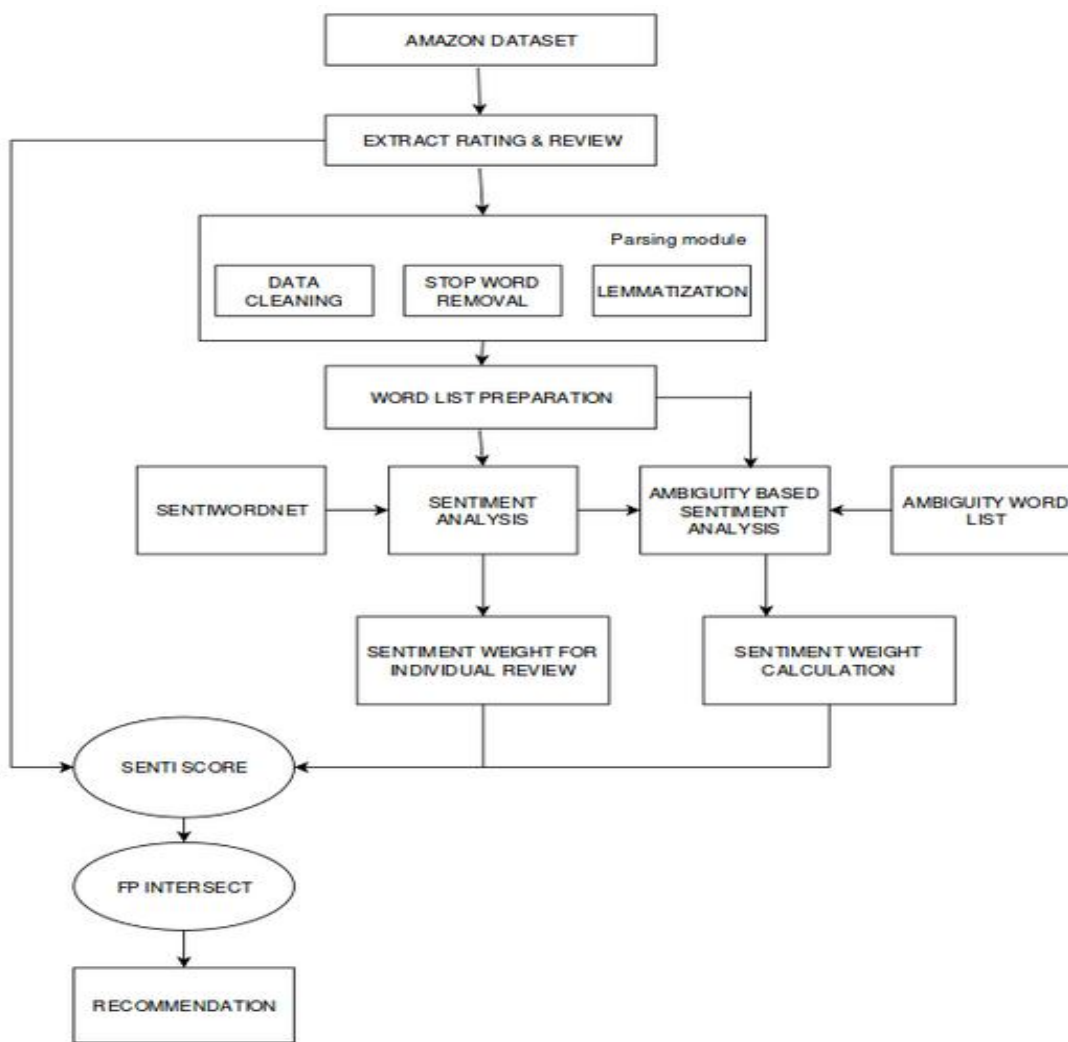


Figure 2: Proposed Solution

## V. CONCLUSION

The complete work concludes that the product recommendation system using sentimental analysis on the basis of FP Intersect algorithm performs better on the basis of rating and ranking of user and also on the basis of user interest. Proposed solution will not only extract the users sentiment but will also help to resolve the issue of ambiguity during sentiment analysis. This research work will help to improve accuracy of sentiment analysis for Amazon e-commerce portals.

## REFERENCES

- [1] Ashraf Elnagar, "Investigation on Sentiment Analysis for Arabic Reviews" published in 13th International Conference of Computer Systems and Applications (AICCSA), 2016, Agadir, Morocco
- [2] Mir Riyanul Islam, "Numeric Rating of Apps on Google Play Store by Sentiment Analysis on User Reviews" published in International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT) 2014
- [3] A.K. Singh, A. Kumar, and A. K. Maurya, "Association rule mining for web usage data to improve websites," in Advances in Engineering and Technology Research (ICAETR), 2014 International Conference on. IEEE, 2014, pp. 1–6.
- [4] S. Rao and P. Gupta, "Implementing improved algorithm over Apriori data mining association rule algorithm 1," 2012.
- [5] O. R. Zaiane, "Building a recommender agent for e-learning systems," in Computers in Education, 2002. Proceedings. International Conference on. IEEE, 2002, pp. 55–59.
- [6] R. Agrawal, R. Srikant et al., "Fast algorithms for mining association rules," in Proc. 20th int. conf. very large data bases, VLDB, vol. 1215, 1994, pp. 487–499.
- [7] P. Nagarnaik and A. Thomas, "Survey on recommendation system methods," in Electronics and Communication Systems (ICECS), 2015 pp. 1496–1501.
- [8] P Devika, R C Jisha and G P Sajeev, "A Novel Approach for Book Recommendation Systems," in International Conference on Computational Intelligence and Computing Research, 2016 IEEE.
- [9] K. Hong, H. Jeon, and C. Jeon, "Userprofile-based personalized research paper recommendation system," in Computing and Networking Technology (ICCNT), 2012 8th International Conference on. IEEE, 2012, pp. 134–138.
- [10] J. Yang, Z. Li, W. Xiang, and L. Xiao, "An improved apriori algorithm based on features," in Computational Intelligence and Security (CIS), 2013 9th International Conference on. IEEE, 2013, pp. 125–128.
- [11] Y. W. Lo and V. Potdar, "A review of opinion mining and sentiment classification framework in social networks," in 3rd IEEE International Conference on Digital Ecosystems and Technologies. 2009, pp. 396–401.
- [12] P. Jomsri, "Book recommendation system for digital library based on user profiles by using association rule," in Innovative Computing Technology (INTECH), 2014 Fourth International Conference on. IEEE, 2014, pp. 130–134.
- [13] M. Al-Maolegi and B. Arkok, "An improved apriori algorithm for association rules," arXiv preprint arXiv:1403.3948, 2014.
- [14] F. Coenen, P. Leng, and S. Ahmed, "Data structure for association rule mining: T-trees and p-trees," IEEE Transactions on Knowledge and Data Engineering, vol. 16, no. 6, pp. 774–778, 2004.
- [15] P. Ramya and G. Sajeev, "Building web personalization system with time-driven web usage mining," in Proceedings of the Third International Symposium on Women in Computing and Informatics ACM, 2015, pp. 38–43.
- [16] G. Sajeev and P. Ramya, "Effective web personalization system based on time and semantic relatedness," in Advances in Computing, Communications and Informatics (ICACCI), 2016 International Conference on. IEEE, 2016, pp. 1390–1396.