Implementation of Product Recommendation Model Using Sentiment Analysis

Ruchira Muchhal[1] & Mohit Jain[2]

[1]MTech Scholar- CSE Department, BM College of Technology, Indore, India.
[2]Assistant Professor - CSE Department, BM College of Technology, Indore, India.

ABSTRACT   *Descriptive data mining is always useful for users to determine most closed recommendations according to their behaviour and interest. Different studies are undergoing in dimension of product recommendation for e-commerce portals and improve recommendation process with minimum time and find that most of the recommendation systems only work on 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 online shopping website. An approach is used and implementation is 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 is implemented on need of opinion mining with respect to E-commerce portal and other websites.*

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

I. INTRODUCTION

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.

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 Flipkart, Amazon, and 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 analyse users view point or user feedback to analyse 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 analyse about organizations 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.

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template

measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

II. LITERATURE REVIEW

A. Existing Work

Before P Devika et al. In [1] Proposed recommendation algorithm using FP Intersect 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. 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.
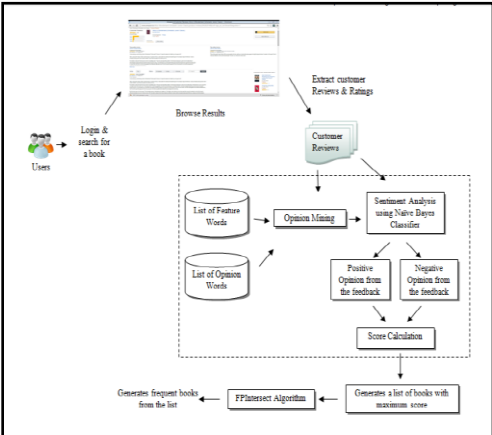

Fig 1. Existing solution by Devika [8]

B. RELATED WORK

A relevant study has been made to analyze problems among existing work. A brief of comparison is shown in Table 1.

TABLE 1. COMPARITIVE TABLE

AUTHOR	TITLE	SUMMARY
P. Jomsri [2]	Book recommendation system for digital library based on user profiles by using association rule.	Used pattern mining algorithm to mine data from the database and generate association relationship.
J. Yang, Z. Li, W. Xiang, and L. Xiao [3]	An improved Apriori algorithm based on features.	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.
P. Nagarnaik, A. Thomas [4]	Survey on recommendation system methods.	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.
Ashraf Elnagar [5]	Investigation on Sentiment Analysis for Arabic Reviews	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.

III. PROBLEM STATEMENT

In e-commerce portal, product rating and review is the important aspects for the study of relevant papers. Every individual needs trust and if any he/she is purchasing online then trust becomes vital. There are the sentiments of user which can be analysed and helps in extraction of user opinion with discovering reviews on the basis of product view. Existing solution develops rating and feedback in numeric form with using Arabic language and performs [5] sentimental analysis on it. Consequently, ambiguity, frequent updations, integrations and data analysis are not considered by author in base work.

Product recommendation study is done by base paper authors to explore that recommendation requires information of user interest, text and review. These are helpful in frequently extracting popularity and reviewing product through shopping sites. An improved product recommendation is implemented by author [1] by using FP intersect algorithm concluding Apriori algorithm. FP intersect is used to overcome the drawback of Apriori algorithm for transactional dataset. On multiple batches FP intersect

algorithm is performed and analysis is done on revised information. Sentiment analysis does not include ambiguity analysis, but ambiguity affects review analysis on the basis of user interest. Issues like:

1. They have used small data for this task, so large data cannot be used for this purpose.
2. Sentimental analysis along with ambiguity analysis is also required in the complete work for the extraction of user review.
3. Improved sentiments are required to perform for large data processing.
4. They have not observed any ambiguity factor on the basis of polarity.

IV. METHODOLOGY

Products are identified for the purpose of recommendation on the basis of user interest including popularity factor and similarity. Similarity of different products is checked on the basis of likes and dislikes of users, their ratings and reviews. Complete study concludes the approach for the implementation of product recommendation on the basis of interest for better recommendation.

Experimental analysis of the proposed work is elaborated step by step as:

1. Amazon dataset is used along with review and rating as the input with different category of the product.
2. Sentimental Analysis on the basis of ambiguity is done.
3. Individual pre-processing by using the technique of tokenization and lemmatization for the extraction of keywords in done using Stanford Lemmatize.
4. Sentiwordnet list is used as WordNet dictionary and identification of polarity of individual words.
5. Polarity analysis of individual words and respective review are done to convert text sentiments into numeric sentiments representation.
6. Sentiment weight calculation for a particular product and weight for individual review is estimated through sentiment score of individual user review and rating along with the ambiguity analysis.
7. In the next step Frequency Pattern Intersect Algorithm (FP Intersect) is applied on the calculated score to get better recommendation.

8. Finally, product recommendation will be displayed based on sentiment analysis.

The complete solution has been demonstrated in block diagram which is shown in figure 2;
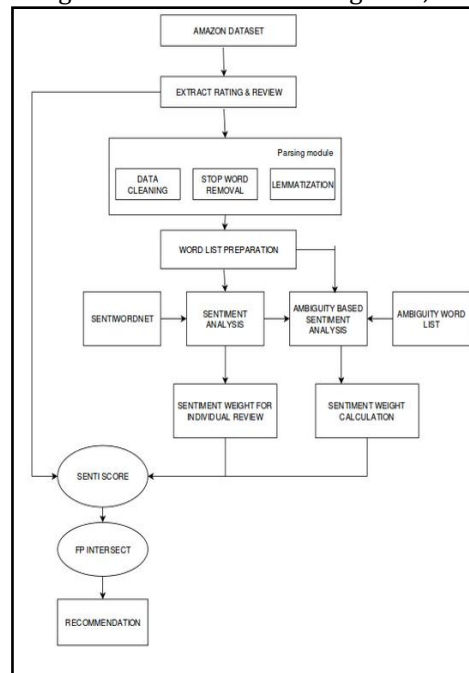

Fig. 2 Proposed solution architecture

V. EXPERIMENTAL ANALYSIS

The complete proposed solution has been implemented for Amazon dataset using Hadoop ecosystem and Java Programming. Here, java based big data application has been implemented to parse different size of data. Amazon dataset has been considered for e-commerce background and different category such as electronics and books has been considered as input for different products. Different size input blocks has been used to evaluate performance on different parameters. Computation time has been considered as performance parameter and evaluated on single node hadoop cluster. Snapshot of proposed implementation has been shown from figure 3-4.
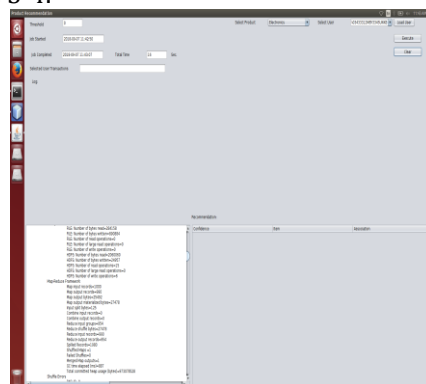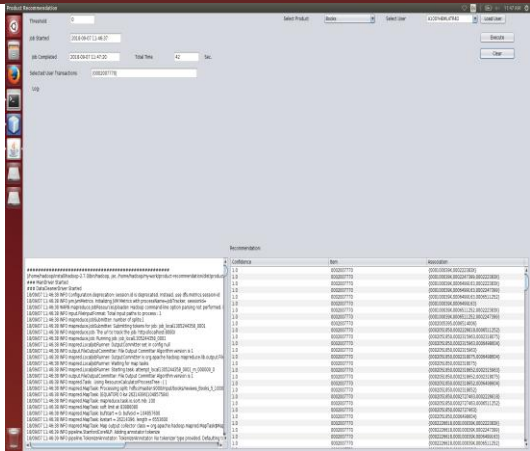

Fig. 3 Home Screen

Fig. 4 Recommendation Screen

The computation time of complete implementation has been evaluated on basis of Single Node Configuration which is shown below; result comparision

Product Category	Computation Time
Book	25 Seconds
Electronics	22 Seconds

Comparison Chart

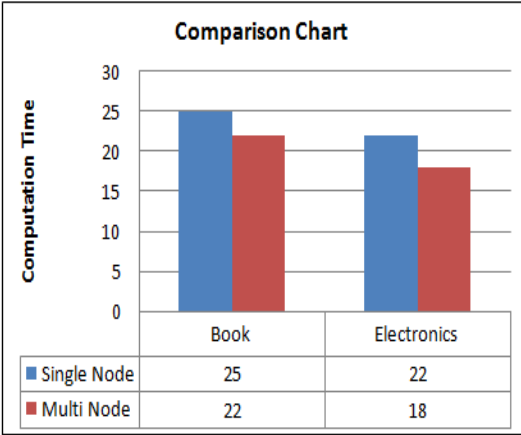|	Book	Electronics
Single Node	25	22
Multi Node	22	18

Fig. 5 Result Comparision Chart

VI. 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 not only extracts the user's sentiment but will also help to resolve the issue of ambiguity during sentiment analysis. This research work helps to improve accuracy of sentiment analysis for Amazon e-commerce portals.

REFERENCES

1. 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.
2. 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.
3. 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.
4. P. Nagarnaik and A. Thomas, "Survey on recommendation system methods," in Electronics and Communication Systems (ICECS), 2015 pp. 1496–1501.
5. Ashraf Elnagar, "Investigation on Sentiment Analysis for Arabic Reviews" published in 13th International Conference of Computer Systems and Applications (AICCSA), 2016, Agadir, Morocco
6. 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.
7. 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
8. 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.
9. S. Rao and P. Gupta, "Implementing improved algorithm over Apriori data mining association rule algorithm 1," 2012.
10. O. R. Za'ıane, "Building a recommender agent for e-learning systems," in Computers in Education, 2002. Proceedings. International Conference on. IEEE, 2002, pp. 55–59.
11. 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.
12. 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.
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.