International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 04 | Apr 2019 www.irjet.net p-ISSN: 2395-0072

Online Sequential Behaviour Analysis using Apriori Algorithm

Gaurav Vishwakarma[1], Shradhey Parte[2], Heem Joshi[3], Jay Patel[4], Pravin Jangid[5]

[1,2,3,4]Dept. of Computer Engineering, Shree L.R. Tiwari College of Engineering
Mira Road (East), Thane- 401107, Maharashtra, INDIA
[5]Assistant Professor, Dept. of Computer Engineering, Shree L.R. Tiwari College of Engineering
Mira Road (East), Thane- 401107, Maharashtra, INDIA

Abstract – *Nowadays a lot of data is available in sequential format. With the emergence of data mining and its application the business sector has been benefited in the form of extraction and prediction algorithms. This has helped to mine such sequential data to form behavioural pattern of such data and make predictions. Recommendation System is one such tool that has been used by almost every E-commerce site. This project explores scope of frequent item set based recommendation by implementing Apriori algorithm which is mainly used to find frequently purchased items/products. The key idea behind this recommendation is that any item set that occurs frequently together must have each item (or any subset) occur at least as frequently.*

***Key Words*: Recommendation System, Apriori Algorithm, Association Rule, Frequent Item set.**

1. INTRODUCTION

Recommendation systems have become extremely common in recent years. In definition, goal of a Recommender System is to generate relevant recommendations to a user for items or different products. Recommendation systems usually produce a list of recommendations in one of two ways - through collaborative filtering or content-based filtering. In Collaborative filtering, it approaches building a model from a user's past activities (items that are previously purchased and/or numerical ratings given to those items) as well as similar decisions made by other users; then use that model to projection items (or ratings for items) that the user may have a concern in.

The most popular recommendation applications in E-commerce are probably books, research articles, search queries, movies, music, news, social tags, and products in general. There are also recommendation systems for life insurance companies, jokes, experts, restaurants, financial services and Twitter followers.

In this work, we are dealing of frequent item set based recommendation using Apriori Algorithm which works on concept of association rules. Example "If a customer purchases shirt then he also buys tie or pants in 70% of the cases". The algorithm searches out frequently purchased items and those items are then suggested as a recommendation to the customer.

2. EXISTING SYSTEM

Today, E-commerce sites use recommendation systems on a large scale to boost their business. The products can be recommended based on the extent of the overall sale with regards to a site, based on the suggestions to the customers, or based upon an analysis of the extra buying behavior of the customer, as a prediction for difficult buying behavior. This methodology is used by retailers all over the world to determine which items are purchased together. Also, they face cold start problem i.e.

1) How to recommend a new user in which case there is no browse history?
2) How to recommend new items which has no purchase history?

It also gives recommendations based on the area of interests of the user, customer searches and also suggests products based on it. For e.g. Amazon or Flip cart uses user view data i.e. if any customer or user searches a product from a specific category the system suggests a product form the same category. Also based on the current search by the user, the site recommends products. Every user who visits the site may not buy a product. They can just go through it and based on those real-time search results the site recommend a product.

3. PROPOSED SYSTEM

Apriori is designed to operate on databases containing transactions and generate association rules, while using a "bottom up" approach, which means that frequent subsets are extended one item at a time and groups of candidates (the candidate set contains all the frequent k-length item sets) which are tested against the data. The algorithm terminates when no further successful extensions are found.

By generating sets of data, we calculate support and confidence of itemsets. We do not calculate support and confidence for itemsets which do not occur together to reduce redundancy of data. The process works in multiple iterations.

First Iteration: An itemset consisting of two products is associated together and support and confidence is calculated for each such itemset. Itemsets are not made for products which do not occur together in transaction since their support and confidence value is zero by default.

Second Iteration: Another item is added to itemsets from previous iteration keeping in mind that the new item must be associate with both of the items from previous itemset thereby creating a new itemset with three items. Again, the support and confidence of the itemset is calculated.

Further Iterations: Depending on the availability of items that can satisfy the same conditions from second iteration, more iterations can exist.

id	iter_	itemset_	support_
27	1	2 17 28	1
29	1	2 17 30	1
62	1	2 28 30	1
69	2	11 14 20	1
71	2	11 14 24	1
73	2	11 14 28	1
74	2	11 14 29	1
87	2	11 19 23	2
95	2	11 20 24	1
97	2	11 20 28	1
98	2	11 20 29	1
108	2	11 24 28	1

Figure 3.1: Itemset Generated in database

4. SYSTEM ANALYSIS

4.1 Front-end layout for the Recommendation:

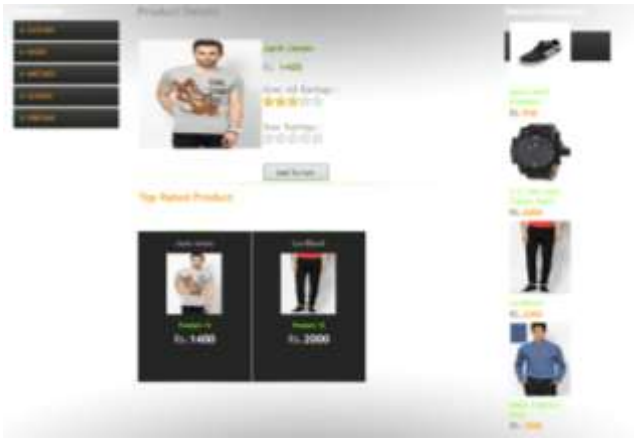

Figure 4.1.1: Recommendation page of the web application

4.2 System Overview Flowchart:

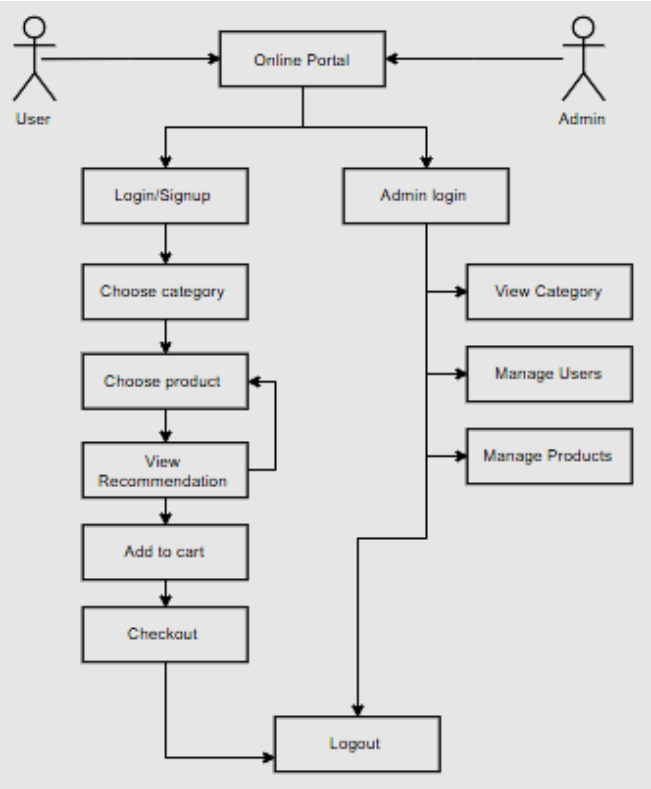

Figure 4.2.1: Flowchart of proposed system

4.3 Proposed System Algorithm:

Apriori Algorithm:

Apriori is designed to operate on databases containing transactions. It applies an iterative approach or level-wise search where k-frequent itemsets are used to find k+1. This algorithm is best suited for recommendation system since such system generally contain a large amount of data that can be grouped together.

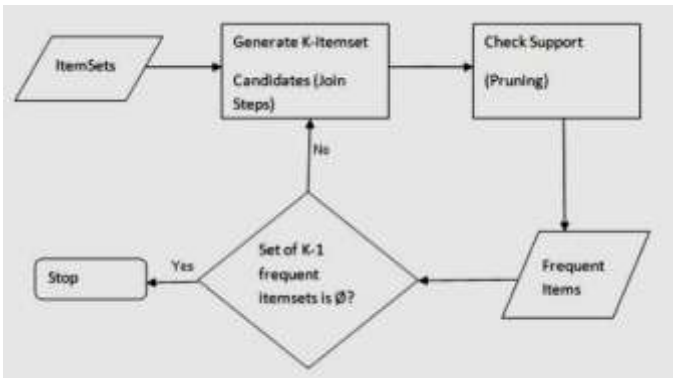

Figure 4.3.1: Flowchart of proposed system algorithm

5. IMPLEMENTATION OF SYSTEM:

The main page consists of Login Portal. After Login the customer can view categories of Products available. The customer can select desired category and proceed to browse the available products. The customer can now check the listed price of the products. The customer will be displayed recommended products on the right side and top-rated products below the main browsing screen.

While browsing the customer can add products to cart based on their requirement. The cart will store the products the customer wants to buy. After this the customer can further browse more products or proceed to checkout.

Figure 5.1: Cart with items

6. APPLICATIONS:

The basic objective of this system is to let user have a fair idea of what a Recommendation system is and how it is useful for an E-commerce system.

7. FUTURE SCOPE:

Recommendation system holds a strong future in E-commerce on a web. Given the current working and design of the proposed systems, there is definitely a place for future enhancements.

In future works, the algorithm can be extended to web content mining, web structure mining, etc. The work can also be extended to extract information from image files.

Lastly, enhancements in the recommendation algorithms used can help increase the accuracy of the system and therefore can help towards the research and development of the topic.

8. CONCLUSIONS:

Recommendation system is a novel interactive technology for fetching additional data for any business from its transaction-oriented database of customers and thus providing different platform of growth to the business. This system helps the customers to find products which they want to buy from the site.

Recommendation system gives benefits to customers by enabling them to find products which they can additionally buy. Conversely, they also help business by generating more sales by advertising more products to customers, increasing their revenue. Recommendation systems are speedily becoming essential tools in E- commerce on the web.

REFERENCES

[1] V. Mayer-Schönberger and K. Cukier, Big Data: A Revolution That Transforms How We Work, Live, and Think. Houghton Mifflin Harcourt, 2012.

[2] Lakhina, M. Crovella, and C. Diot, "Diagnosing network-wide traffic anomalies," in Proceedings of ACM Special Interest Group on Data Communication (SIGCOMM), 2004.

[3] E. E. Papalexakis, A. Beutel, and P. Steenkiste, "Network anomaly detection using co-clustering," in Proceedings of IEEE/ACM International Conference on Advances in Social Network Analysis and Mining (ASONAM), 2012.

[4] Z. Guo, Z. Li, and H. Tu, "Sina microblog: an information-driven online social network," in 2011 International Conference on Cyberworlds (CW), 2011.

[5] M. Bishop, Pattern recognition and machine learning. springer, 2006.

[6] J. D. Lafferty, A. McCallum, and F. C. N. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," in Proceedings of ACM International Conference on Machine Learning (ICML), 2001.

[7] J. Bertsimasi, A. J. Mersereauy, and N. R. Patel, "Dynamic classification of online customers," in Proceedings of SIAM International Conference on Data Mining (ICDM), 2003.