

Recommendation System using Apriori Algorithm

Karandeep Singh Talwar¹ Abhishek Oraganti² Ninad Mahajan³ Pravin Narsale⁴

^{1,2,3,4}Department of Information Technology

^{1,2,3,4}Mumbai University Padmabhushan Vasantdada Patil Pratishthan's College of Engineering, Sion-Mumbai-22

Abstract— Recommendation Systems are changing from novelties which were used by a few E-commerce sites, to tools that are almost-shaping the world of E-commerce. Many of the largest commerce web sites are already using recommendation systems to encourage their customers, locate products. A recommendation system learns from a customer and recommends products that he/she will investigate most necessary from the available products. This research work is to recommend a product using Apriori algorithm. This algorithm is mainly used to find frequently purchased items/products. Its aim is to detect association rules.

Key words: Recommendation System, Apriori, Web Usage Mining, Association Rule

I. INTRODUCTION

Recommendation Systems are used to suggest products for E-commerce sites. A recommendation system is a software evolved from a new class of data analysis which applies knowledge discovery techniques to the misfortune of making product recommendations during live customer dealings. These systems are achieving widespread attainment in E-commerce nowadays, especially after the advent of the Internet. The tremendous accretion of customers and products poses three key challenges for recommendation systems in the E-commerce domain. These are: producing high standards of recommendations, display many recommendations per second for millions of customers and products, and achieving maximum coverage in the turn of data sparsity. One affluent recommendation system technology is collaborative filtering, which works by matching customer preferences to other customers while making recommendations.

This paper presents a new recommendation technique using Apriori algorithm. Its aim is to detect association rules. Example "If a customer purchases bread then he also buys butter or jam in 70% of the cases". The algorithm searches out frequently purchased items and those items are then suggested as a recommendation to the customer.

II. 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. Broadly, these techniques are manifestations of personalization which the site offers its customers. 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?

2) How to recommend a new items?

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 search results the site recommends a product.

III. PROPOSED SYSTEM

In our proposed system, we are planning to use a methodology that leverages the transaction history. As Apriori is designed to operate on databases containing transactions and generate association rules, while using a "bottom up" approach, where 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. The architectural design of the proposed system is shown in Figure A. This methodology eliminates the first problem which is making recommendations for new customer. As this proposed system is using transaction oriented data, it won't require the customer's profile to recommend products.

The Recommendation System using Apriori algorithm can be implemented with four main functionalities: User Interface Component, Data Extraction, Web Usage Mining and Pattern Recognition.

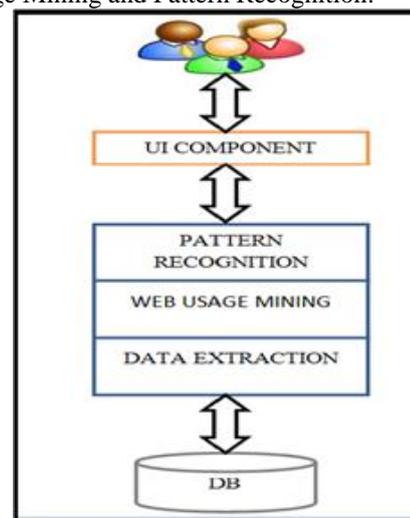


Fig. 1: Architecture of Recommendation System

A. User Interface Component:

This module communicates between users and the web usage mining system. It allows the user to interact with the system by specifying a web usage mining query or task, providing information to help focus the search, and

performing exploratory web usage mining based on the intermediate web usage mining results.

B. Data Extraction:

Data extraction is the act or process of retrieving data (usually unstructured or poorly structured data) out of data sources for further data processing or data storage (data migration). In our proposed system we are extracting data from database on which we are applying the Apriori algorithm.

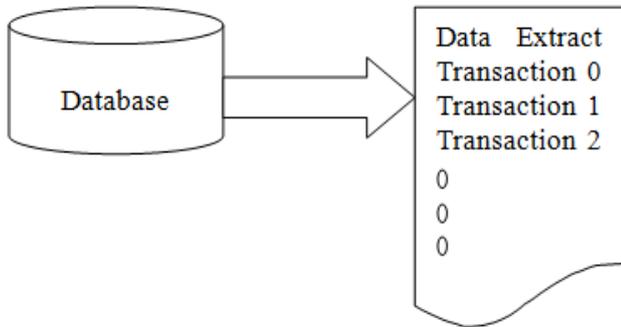


Fig. 2: Data Extraction

C. Web Usage Mining:

Web usage mining is achieved first by reporting visitors traffic information based on web server log files and other source of traffic data. Web log file is one way to collect Web traffic data. After the Web traffic data is obtained, it may be combined with other relational databases, over which the data mining techniques are implemented through some data mining techniques such as association rules, clustering etc.

D. Pattern Recognition:

Pattern Recognition or evaluation is used to identify the truly interesting patterns representing knowledge based on some interestingness measures. This component typically employs interestingness measures and interacts with the Web Usage Mining modules so as to focus the search toward interesting patterns. Recommendation System should be able to display the discovered patterns in multiple forms, such as rule, tables, pie or bar charts, cubes or other visual representation. In our Recommendation systems we provide recommendation of product through product images. It may use interestingness thresholds to filter out discovered patterns. Alternatively, the pattern evaluation module may be integrated with the mining module, depending on the implementation of the data mining method used.

IV. WORKING

Working or flow of the recommendation system after user logs in to the system is depicted in Figure C (i.e. activity diagram of system).

V. ILLUSTRATION OF ALGORITHM

Apriori algorithm is a classic algorithm for learning association rules. We are applying Apriori on a database that contains the transaction (e.g. a collection of items purchased by customers etc.)

Apriori algorithm is easy to execute and very simple. It is used to mine all frequent item sets in database. The algorithm makes many searches in database to find

frequent item sets where k-item sets are used to generate k+1-itemsets. Each k-item set must be greater than or equal to minimum support threshold frequency. Otherwise, it is called candidate item sets. In the first, the algorithm scans the database to find frequency of 1-itemsets that contains only one item by counting each item in the database. The frequency of 1-itemsets is used to find the item sets in 2-itemsets which in turn is used to find 3-itemsets and so on until there are not any more k-item sets. If an item set is not frequent, any large subset from it is also non-frequent. In this condition pruning from the search space in database is conducted. Figure D illustrates the flow of Apriori Algorithm.

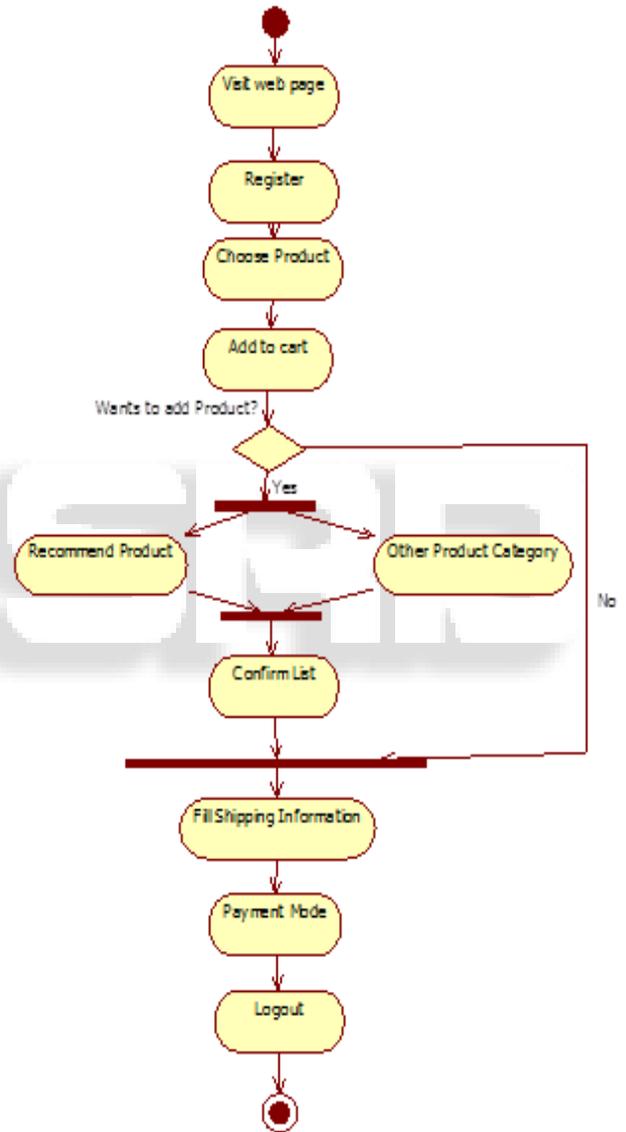


Fig. 3: Activity Diagram

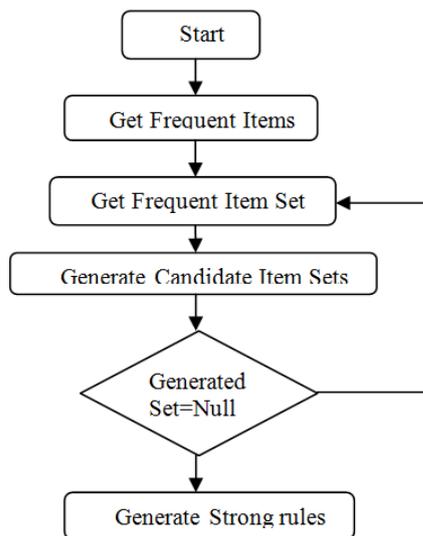


Fig. 4: Flow Chart of Apriori algorithm

VI. CONCLUSION

Recommendation system is a novel interactive technology for fetching additional data for any business from its transaction oriented database of customers. 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 must buy. Conversely, they also help business by generating more sales, increasing their revenue. Recommendation systems are speedily becoming essential tools in E-commerce on the web.

VII. FUTURE SCOPE

Recommendation system holds a strong future in E-commerce on a web. With years to come this recommendation system shall be utilized at various sites to recommend various services also. In future 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.

REFERENCES

- [1] <http://www.recommenderbook.net/teaching-material/slides>
- [2] Recommender Systems Handbook, Ricci, F.; Rokach, L.; Shapira, B.; Kantor, P.B. (Eds.), 2011, Springer.
- [3] <http://ijana.in/papers/6.11.pdf>
- [4] <http://www.win.tue.nl/~laroyo/2L340/resources/recommender-systems-e-commerce.pdf>
- [5] Data Mining Concepts and Techniques 2nd Ed By Kamber pages 234-242.