A Product Recommendation Model Prototype

Nikeeta Brijwasi, Shefali Jain, Sudarshana Singh, Vijaya Rani, Matthew A. Lanham
Purdue University Krannert School of Management
nbrijwas@purdue.edu; jain291@purdue.edu; singh554@purdue.edu; vrani@purdue.edu; lanhamm@purdue.edu

DAWN OR DOOM 2018 — EMERGING TECHNOLOGY – RISKS AND REWARDS — Two-Day Conference, Nov. 5 & 6

BIAC — Business Information and Analytics Center

Abstract

Personalized recommendations are an important part of many e-commerce applications. Recommender systems are gaining tremendous popularity across companies, such as Amazon and Netflix due to their effectiveness and efficiency in helping users filter through enormous numbers of items and in helping enterprisers increase their sales. This working paper presents an e-commerce based recommendation system that can be used by the retailer to increase profits and differentiate itself from its competitors. The model is based on collaborative filtering and creating clusters among the customers. Recency-Frequency-Monetary Value (RFM) and Market Basket Analysis (MBA). The model trains itself using past purchases to predict future recommendations for customers of similar nature.

Introduction

Retailers are traditionally interested in understanding the composition of their customers' market baskets, since valuable insights for designing micro-marketing and/or targeted cross-selling programs can be derived. Recommender systems are technologies that assist business to implement such strategies. It can help the retailers determine how they can increase the probability to cross-sell, establish customer loyalty, decrease drop-off rates, increase revenue by targeted advertisements, improving customer satisfaction and retention, efficiently managing inventory.

Research Question:

- Can we develop a product recommendation system that makes business sense using popular algorithms as well a decision-support tool prototype?

Literature Review

Various approaches have been studied to generate reliable and scalable product recommendation systems. These are the studies we reviewed and how our approach relates to them.

Sr No	Paper	MBA	RFM	Reg	CF	RF
1	Prediction Approach			●	●	
2	Higher Education Recommendation			●		
3	Predicting Online Behavior			●		
4	Random Forest for Online Behavior					●
5	Clustering item preference		●			
6	Building Prediction Model	●				
7	Product Recommendation		●		●	
8	Recommendation for grocery store				●	
9	Cross category dependence	●			●	
10	Market Basket Analysis	●				
	Product Recommendation Model	●	●			

Methodology

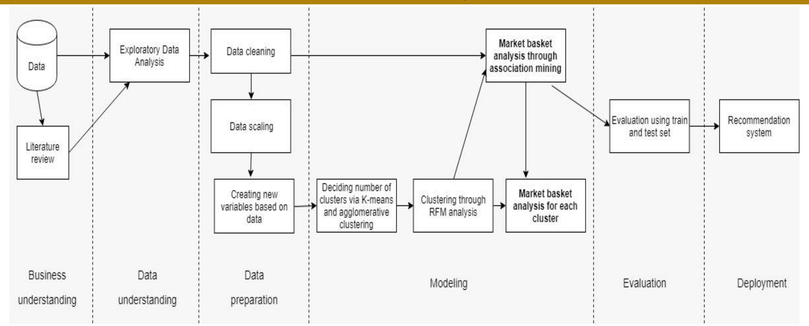


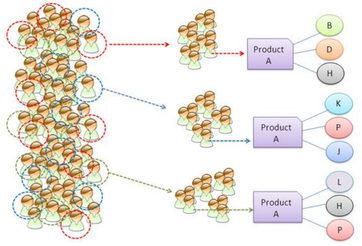
Figure 2. Methodology

Data Cleaning & Pre-Processing

We categorized various products to reduce the number of variables. This helped improve the model in terms of different recommendations.

Methodology (Approach) Selection

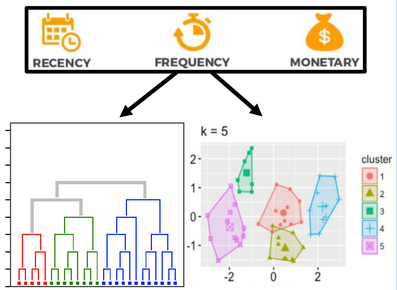
Collaborative filtering: recommends items by finding similar users to the active user (to whom we are trying to recommend an object). A specific application of this is the user-based Nearest Neighbor algorithm.



Transaction 1	
Transaction 2	
Transaction 3	
Transaction 4	
Transaction 5	
Transaction 6	
Transaction 7	
Transaction 8	

Association rule mining: used to identify relationships among a set of items in a database. These relationships are based on the co-occurence of the data items. How items are purchased together allow one to easily create rules that can help in product placement, promotional strategies, and in this case recommended items.

Hierarchical clustering: Customers were divided into clusters basis their RFM (Recency, Frequency, Monetary) signals. Hierarchical agglomerative clustering is used which is basically a bottom-up algorithm which treat each document as a singleton cluster at the outset and then successively merges (or *agglomerates*) pairs of clusters until all clusters have been merged into a single cluster.



Non hierarchical clustering: K-means clustering was performed on test and train dataset, doing a 50/50 split of the customer data based on RFM variables. Based on the elbow plot curve 5 was chosen as the number of clusters.

Results

The K-means clustering and agglomerative clustering techniques both suggested that we should split customers into these 5 groups.

Cluster Name	Cluster Description
Upscale Buyers	Buy most expensive items (high price affinity), but infrequent, low vertical diversity
On the brink customers	Inactive customers with lowest spend and have the highest probability of lapsing
Average Joe	Active buyers with medium frequency and spend
Bottom of the Barrel	Infrequent single time buyers with lowest GMV and transaction frequency, no high value purchase
Top of the Hill	Most frequent and active transactors with highest revenue contributing cluster

Figure 3. Clustering

The accuracy of market basket analysis (association rules) model as well as RFM analysis was quite high for our model. Thus we decided to use both the models for our recommendation system.

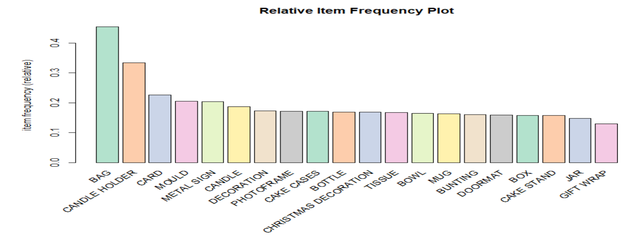


Figure 4. Market Basket Analysis

Conclusions

Today consumers are exposed to an increasing variety of products and information never seen before. This analysis can be used in managing the product placement on shelves in stores or identify potential products online. Furthermore, this method can prove to fetch more profit to the seller. A good recommendation systems can fulfill customer needs and expectations while help maintain or improve customer loyalty. Also help in cross-selling, and increase revenue by targeted advertisements.

We found that our solution of RFM (Recency-Frequency-Monetary) based clustering & user-based collaborative filtering could be useful for retailers where they can predict the customer needs based on only their past purchase patterns. No more information is required to understand the item content than what the customer has purchased previously.

Our solution also would make business sense and thus increase the chance that decision-makers would actually use it.

Decision Support Tool Prototype

Video of our R-Shiny app prototype here: **https://goo.gl/vj4t72**