Alexa! Let's Order Food...

Mayank Sharma, Rashmi Ojha, Srishti Agarwal, Ronak Chaudhary

Department of Computer Science & Engineering, Texas A&M University

Introduction

E-commerce websites allow users to rate the products or services after purchasing which further serve as suggestions to other users. These feedbacks are influential to people's decisions on whether to buy the product or not. The reviews and ratings are also an important way to understand the behavior and buying pattern of users for future growth of companies.

Motivation

Based on the Amazon Dataset, we wanted to build a category specific recommendation system so that users looking for certain category of Food Products like 'beverages' or 'Pet Food', may explore relevant products directly with ease. We compared different methods to provide a good recommendation and tried a new category based item-item filtering approach, which was unfortunately not successful. This project is particularly important for the following reasons:

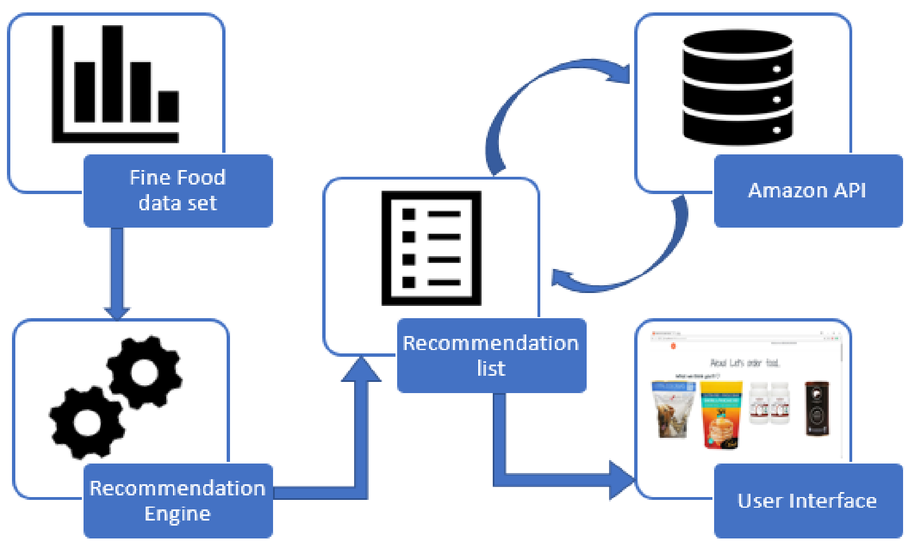
1) Business aspect (Boost a company's revenue)
2) Learning aspect

Dataset

We have taken the Amazon Fine Food Reviews Dataset to build our recommendation engine. It consists of 568,454 food reviews which Amazon users left from October, 1999 to October, 2012 encompassing 256,059 users and 74,258 food products. Reviews include product and user information, ratings, and a plain text review. It also includes some reviews from other Amazon categories. Each review has values in the following format:

Product/**productId**:	B001E4KFG0
Review/**userId**:	A3SGXH7AUHU8GW
Review/**profileName**:	delmartian
Review/**helpfulness**:	1/1
Review/**score**:	5.0
Review/**time**:	1303862400
Review/**summary**:	Good Quality Dog Food
Review/**text**:	I have bought several of the Vitality canned dog food products and have found them all to be of good quality. The product looks more like a stew than a processed meat and it smells better. My Labrador is finicky and she appreciates this product better than most.

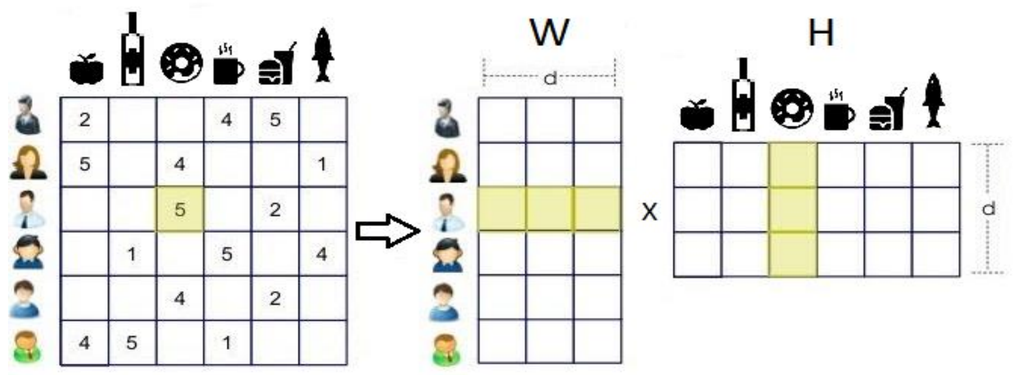
Architecture



Approaches

Our system implements a Food Recommender Engine providing user-specific / personalized recommendations along with generic or popular item recommendations. We implemented two well-known algorithms and tried to develop a new category specific item recommendation algorithm. We used Matrix Factorization and item-item based Collaborating Filtering for developing the initial system and chose Matrix Factorization for the final application.

1. Matrix Factorization



2. Item-Item Collaborative Filtering

```
Input: set of reviews
Output: item-item similarity matrix
1. For each item in product catalog, I1
2.     For each customer C who purchased I1
3.         For each item I2 purchased by customer C
4.             Record that a customer purchased I1 and I2
5.     For each item I2
6.         Compute the similarity between I1 and I2
```

3. Category specific Item Recommendation

Our idea was to first categorize all items into certain categories based on review text and summary. Then recommend items based on user's selection of a particular category. We were not able to successfully implement the category specific recommendation engine as the clusters obtained using Review Texts were not a good representation of the requisite categories.

Key Takeaways

1. Matrix Factorization is better than User-User and Item-Item Collaborative Filtering approaches.
2. Clustering of items based on review text is very hard to obtain.

Negative Impact

1. Shilling attacks by cheap and junk food companies can promote unhealthy eating habits in society.
2. Popular products recommended by Item-Item CF approach may not suit everyone's taste and budget.

References

1. J. McAuley and J. Leskovec. *From amateurs to connoisseurs: modeling the evolution of user expertise through online reviews.* WWW, 2013.
2. Greg Linden, Brent Smith, and Jeremy York. 2003. *Amazon.com Recommendations: Item-to-Item Collaborative Filtering*. IEEE Internet Computing 7, 1 (January 2003), 76-80.