OUTLINE

- INTRODUCTION
- DIFFERENT TYPES OF FILTERING SYSTEMS
- RECOMMENDATION ALGORITHM CHALLENGES
- USER AND ITEM BASED FILTERING
- TRADITIONAL COLLABRATIVE FILTERING
- CLUSTER MODELS
- SEARCH BASED METHODS
- ITEM-TO-ITEM COLLABRATIVE FILTERING
- SCALABILITY
- CONCLUSION
Recommendation algorithms are best known for their use on e-commerce websites where they use input about a customer's interests to generate a list of recommendation items.

Many applications use only the items that customers purchase and explicitly rate to represent their interests, but they can also use other attributes, including items viewed, demographic data, subject interests, and favorite artists.

35% of Amazon revenue is from recommendation engine.
DIFFERENT TYPES OF FILTERING SYSTEMS

1. Collaborative filtering.

2. Cluster models.


4. Demographic based Recommender system.

5. Knowledge based Recommender system.

6. Hybrid filtering.
A large retailer might have huge amounts of data, tens of millions of customers and millions of distinct catalog items.

Many applications require the results set to be returned in real time, in more than half a second while still producing high quality recommendations.

New customers typically have extremely limited information based on only a few purchases or product ratings.

Older customer can have a glut of information based on thousands of purchases and ratings.

Customer data is volatile: each interaction provides valuable customer data, and the algorithm must respond immediately to the new information.

Recommendation Algorithm Challenges
COLLABORATIVE FILTERING AND CLUSTER MODELS FOCUS ON FINDING SIMILAR SET OF CUSTOMERS WHOSE PURCHASED AND RATED ITEMS OVERLAP WITH THE USER’S PURCHASED AND RATED ITEMS.

SEARCH-BASED METHODS AND ITEM-TO-ITEM COLLABORATIVE FILTERING FOCUS ON FINDING SIMILAR ITEMS, NOT SIMILAR CUSTOMERS.

USER BASED AND ITEM BASED FILTERING
The algorithm generates recommendations based on few customers who are most similar to the user.

Here the similarity between the two customers is calculated using the cosine similarity.

The component vector are positive for purchased or positively rated items and negative for negatively rated items.

\[
\text{similarity}(\vec{A}, \vec{B}) = \cos(\vec{A}, \vec{B}) = \frac{\vec{A} \cdot \vec{B}}{\|\vec{A}\| \cdot \|\vec{B}\|}
\]
Cosine Similarity

\[ \text{sim}(A, B) = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} \]
LIMITATIONS

- Using collaborative filtering is computationally expensive
- Worst case complexity is $O(MN)$ [M- Number of customers, N- Number. of product catalog].
- The final performance of the algorithm is $O(M+N)$.
- The algorithm encounters severe performance and scaling issues.
- We can reduce M and N by a large factor but with the cost of recommendation quality.
Divide the customer base into many segments and assign the user to the segment containing the most similar customers.

Segments are created using clustering.

The algorithm’s goal is to assign the user to the segment containing the most similar customers. It then uses the purchases and ratings of the customers in the segment to generate recommendations.

Cluster models have better online scalability and performance than collaborative filtering because they compare the user to a controlled number of segments rather than the entire customer base.

CLUSTER MODELS
Cluster models group numerous customers together in a segment, match a user to a segment, and then consider all customers in the segment similar customers for the purpose of making recommendations.

Because the similar customers that the cluster models find are not the most similar customers, the recommendations they produce are less relevant.

online user–segment classification becomes almost as expensive as finding similar customers using collaborative filtering.
SEARCH-BASED METHODS

- Search- or content-based methods treat the recommendations problem as a search for related items.
- The algorithm constructs a search query to find other popular items by the same author, artist, or director, or with similar keywords or subjects.
- If the user has few purchases or ratings, search based recommendation algorithms scale and perform well.
LIMITATIONS

- For users with thousands of purchases, however, it’s impractical to base a query on all the items.
- The algorithm must use a subset or summary of the data, reducing quality.
- In all cases, recommendation quality is relatively poor.
- Recommendations should help a customer find and discover new, relevant, and interesting items.
ITEM-TO-ITEM COLLABORATIVE FILTERING

- Rather than matching the user to similar customers, item-to-item collaborative filtering matches each of the user’s purchased and rated items to similar items, then combines those similar items into a recommendation list.
- To determine the most-similar match for a given item, the algorithm builds a similar-items table by finding items that customers tend to purchase together.
- We could build a product-to-product matrix by iterating through all item pairs and computing a similarity metric for each pair. However, many product pairs have no common customers, and thus the approach is inefficient in terms of processing time and memory usage.
For each item in product catalog, $I_1$

For each customer $C$ who purchased $I_1$

For each item $I_2$ purchased by customer $C$

Record that a customer purchased $I_1$ and $I_2$

For each item $I_2$

Compute the similarity between $I_1$ and $I_2$
Item-based filtering

User 1

User 2

User 3

Purchases

Recommended to User 3
HOW IT WORKS

- It’s possible to compute the similarity between two items in various ways, but a common method is to use the cosine measure, in which each vector corresponds to an item rather than a customer, and the vector’s M dimensions correspond to customers who have purchased that item.

- This offline computation of the similar-items table is extremely time intensive, with $O(N^2 M)$ as worst case, but in real time it is $O(NM)$.

- Given a similar-items table, the algorithm finds items similar to each of the user’s purchases and ratings, aggregates those items, and then recommends the most popular or correlated items. This computation is very quick, depending only on the number of items the user purchased or rated.
SCALABILITY

- Traditional collaborative filtering does little or no offline computation, and its online computation scales with the number of customers and catalog items.

- Cluster models can perform much of the computation offline, but recommendation quality is relatively poor. To improve it, it’s possible to increase the number of segments, but this makes the online user-segment classification expensive.

- Search-based models build keyword, category, and author indexes offline, but fail to provide recommendations with interesting, targeted titles. They also scale poorly for customers with numerous purchases and ratings.

- For item-to-item it is dependent only on how many items the user has purchased or rated. Thus, the algorithm is fast even for extremely large data sets. Because the algorithm recommends highly correlated similar items, recommendation quality is excellent.
1) "RECOMMENDED FOR YOU, THOMAS"

Recommended for you, Thomas

Nayoya Gymnastic Rings for Full Body Strength and Crossfit Training
- $29.87 + $1.274 Prime

Champion Sports NCAA NFHS Certified Lacrosse Ball
- $6.95 - $11.19

Rumble Roller Beastie Hook - Hand Held Massage Tool - Use With RumbleRoller Beastie...
- $13.84 + $1.195 Prime

Muscle Roller Ball Set.: Massage Balls for Deep Tissue, Trigger Point & Myofascial Release...
- $32.97 + $5.97 Prime

Foam Roller, LuxFit Premium High Density Foam Roller - Extra Firm With 1 Year Warranty
- $8.73 - $24.95

2 X LACROSSE BALLS FOR TRIGGER POINT MASSAGE-Fine Toneibil plus... 
- $10.75 + $3.35 Prime

The Dead Wedge - Deadlift Jack Alternative for Your Gym Bag - Raise loaded barbell & plates...
- 4.5 out of 5 stars, $17.95 + Prime

Departments
- Browsing History
- Thomas's Amazon.com
- Today's Deals
- Gift Cards & Registry
- Sell
- Help

Your Browsing History
- View and Edit

Rumble Roller Full Size Extra Firm Black...

Kindle Paperwhite, 6" High Resolution Display...

Rumble Roller Beastie Hook - Hand Held...

Amazon Prime (One Year Membership)

Harney & Sons Black Tea Hot Cinnamon Sunset...

Frequently Bought Together

Total price: $94.90

Add both to Cart

Add both to List

This item: Rumble Roller - Textured Muscle Foam Roller Manipulates Soft Tissue Like A Massage Therapist...
- $69.95

Rumble Roller X-Firm Beastie and Base - Extra Firm Spiky Massage Ball - Comes With Base For...
- $24.95 ($3.12/oz)
Best-selling emerging technology

There is a newer version of this item:

Kindle Paperwhite E-reader, 6" High-Resolution Display (300 ppi) with Built-in Light, Wi-Fi - Includes Special Offers

$119.99

4.8 stars (19,163)

In Stock.

Recommended for You Based on Kindle Paperwhite, 6" High Resolution Display w...
Recommendation algorithms provide an effective form of targeted marketing by creating a personalized shopping experience for each customer.

Unlike other algorithms, item-to-item collaborative filtering is able to meet the challenges.

In the future, we expect the retail industry to more broadly apply recommendation algorithms for targeted marketing, both online and offline.
THANK YOU