

Recommender Systems

Debapriyo Majumdar
Indian Statistical Institute Kolkata

Credits to Bing Liu (UIC) and Angshul Majumdar (IIITD) for some slides

Recommender systems



Business

- How to increase revenue?
- How to recommend items customers like?



Customer

- Too many options.
- How to choose the right one?

Recommender systems



Apple Flip Cover for iPad Mini (Grey)

Write a REVIEW Add to WISHLIST

- Magnetic Connection
- Wake & Sleep Function
- Keyboard Stand
- Face Time

Available with 1 Seller at 700092 [Change](#)

MRP: Rs. 2,999
Rs. 2,299 23% OFF
Selling Price
+ Rs 70 Delivery ?

ADD TO CART

BUY NOW

SOLD BY
[Chillzone-Digital](#) 4.2 / 5

DELIVERED BY ? CASH ON DELIVERY
• Mon, 20th Apr: Rs. 70 ?
Available

10 day Replacement Guarantee. ?

Customers who viewed / bought this product also bought

Since you are looking at this, you may also look at ...


CUSTOMERS WHO VIEWED THIS PRODUCT ALSO VIEWED



Five product recommendations for iPad Mini covers are displayed in a row, each with a star rating and price information.

Product Name	Rating	Price
Apple Book Cover for iPad Mini	★★★★★	Rs 4,999 (40% Off) Rs 2,999
Apple Flip Cover for iPad mini, iPad mini with Retina	★★★★★	Rs 3,000 (16% Off) Rs 2,495
Apple Book Cover for iPad Mini	★★★★★	Rs 4,499 (33% Off) Rs 2,999
Apple Flip Cover for iPad mini with Retina Display	★★★★★	Rs 2,999 (20% Off) Rs 2,399
Apple Book Cover for iPad Mini	★★★★★	Rs 2,900 (31% Off) Rs 1,999

Recommender systems



Pirates of the Caribbean: At World's End (2007)

PG-13 | 169 min | Action, Adventure, Fantasy | 25 May 2007 (USA)

Your rating: ★★★★★★★★ ~/10
Ratings: **7.1**/10 from 404,040 users Metascore: 50/100
Reviews: 1,233 user | 303 critic | 36 from Metacritic.com


Captain Barbossa, Will Turner and Elizabeth Swann must sail off the edge of the map, navigate treachery and betrayal, and make their final alliances for one last decisive battle.

Director: Gore Verbinski
Writers: Ted Elliott, Terry Rossio, 4 more credits »
Stars: Johnny Depp, Orlando Bloom, Keira Knightley | See full cast and crew »

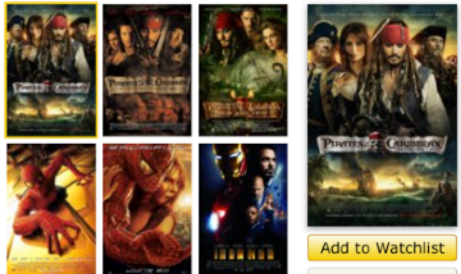
[+ Watchlist](#) [Share...](#)

Nominated for 2 Oscars. Another 20 wins & 36 nominations. [See more awards »](#)

Photos


[130 photos](#) | [385 news articles](#) »

People who liked this also liked... [Learn more](#)



Pirates of the Caribbean: On Stranger Tides (2011)

PG-13 Action | Adventure | Fantasy

★★★★★☆☆ 6.7/10

Jack Sparrow and Barbossa embark on a quest to find the elusive fountain of youth, only to discover that Blackbeard and his daughter are after it too.

[Add to Watchlist](#)

[Next »](#)

Director: Rob Marshall
Stars: Johnny Depp, Penélope Cruz,...

◀ Prev 6 Next ▶

Viewers who liked this movie also liked the other movies

Since you are looking at this page, you may also like...

The Recommendation Problem

- We have a set of users U and a set of items S to be recommended to the users.
- Let p be an utility function that measures the usefulness of item s ($\in S$) to user u ($\in U$), i.e.,
$$p : U \times S \rightarrow R, \text{ where } R \text{ is a totally ordered set (e.g., non-negative integers or real numbers in a range)}$$
- Objective
 - Learn p based on the past data
 - Use p to predict the utility value of each item s ($\in S$) to each user u ($\in U$)

Two main formulations

- Rating prediction: predict the rating score that a user is likely to give to an item that (s)he has not seen or used before
 - Rating on an unseen movie
 - In this case, the utility of item s to user u is the rating given to s by u
- Item prediction: predict a ranked list of items that a user is likely to buy or use

Approaches

Content-based recommendations:

- The user will be recommended items similar to the ones the user preferred in the past

Collaborative filtering (or collaborative recommendations):

- The user will be recommended items that people with similar tastes and preferences liked in the past

Hybrids: Combine collaborative and content-based methods

Content based recommendation

- Will user u like item s ?
- Look at items similar to s ; does u like them?
 - Similarity based on content
 - Example: a movie represented based on features as specific actors, director, genre, subject matter, etc
- The user's interest or preference is also represented by the same set of features (the user profile)
- Candidate item s is compared with the user profile of u in the same feature space
- Determine if u would like s , or
- Top k similar items are recommended

Collaborative filtering

- Collaborative filtering (CF): more studied and widely used recommendation approach in practice
 - k -nearest neighbor
 - association rules based prediction
 - matrix factorization
- Key characteristic: predicts the utility of items for a user based on the items previously rated by other like-minded users (thus, *collaborative*)

k nearest neighbor approach

- No model building
- Utilizes the entire user-item database to generate predictions directly, i.e., there is no model building.
- This approach includes both
 - User-based methods
 - Item-based methods

User based kNN CF

- Let the record (or profile) of the target user be u (represented as a vector), and the record of another user be v .
- The similarity between the target user, u , and a neighbor, v , can be calculated using the **Pearson's correlation coefficient**:

$$\text{sim}(u, v) = \frac{\sum_{i \in S} (r_{u,i} - \bar{r}_u)(r_{v,i} - \bar{r}_v)}{\sqrt{\sum_{i \in S} (r_{u,i} - \bar{r}_u)^2} \sqrt{\sum_{i \in S} (r_{v,i} - \bar{r}_v)^2}}$$

and compute V , is the set of k similar users based on this similarity

- Compute the rating prediction of item i for target user u

$$p(u, i) = \bar{r}_u + \frac{\sum_{v \in V} \text{sim}(u, v) \times (r_{v,i} - \bar{r}_v)}{\sum_{v \in V} |\text{sim}(u, v)|}$$

Problems with user based CF

- The problem with the user-based formulation of collaborative filtering is the lack of scalability:
 - it requires the real-time comparison of the target user to all user records in order to generate predictions
- A variation of this approach that remedies this problem is called item-based CF

Item-based CF

- The item-based approach works by comparing items based on their pattern of ratings across users. The similarity of items i and j is computed as follows:

$$\text{sim}(i, j) = \frac{\sum_{u \in U} (r_{u,i} - \bar{r}_i)(r_{u,j} - \bar{r}_j)}{\sqrt{\sum_{i \in U} (r_{u,i} - \bar{r}_i)^2} \sqrt{\sum_{i \in U} (r_{u,j} - \bar{r}_j)^2}}$$

- After computing the similarity between items we select a set of k most similar items and generate a predicted value of user u 's rating

$$p(u, i) = \frac{\sum_{j \in J} r_{u,j} \times \text{sim}(i, j)}{\sum_{j \in J} |\text{sim}(i, j)|}$$

where J is the set of k similar items

Association rule-based CF

- Transaction database: users, items
 - User \rightarrow item: viewed, bought, liked
- Find association rules such as
 - Bought X , bought $Y \rightarrow$ Bought Z
 - Confidence and support (how strong is this association)
- Rank items based on measures such as confidence, subject to some minimum support
- Further reading: association rule mining