Recommendation Systems - Part 1

Topics covered so far

- Recommendation Systems
    - Introduction to the Recommendation Systems, specific metrics to measure the performance, sparsity of data, and time varying data
    - Examples of datasets
    - Modelling process and simple solutions
    - Improving solutions, Clustering
    - Collaborative Filtering
    - Singular Value Thresholding

Discussion questions

1. Why are recommendation systems useful? Give some examples of recommendation systems.
2. How would you define popularity based recommendation systems?
3. What are the measures to find similarity among users/items in recommendation systems?
4. What are collaborative filtering based recommendation systems and its types?
5. What is matrix factorization technique that is used for recommendation systems?

Recommendation Systems

Recommender systems aim to predict users' interests and recommend product items that quite likely are interesting for them.

Why are they useful?

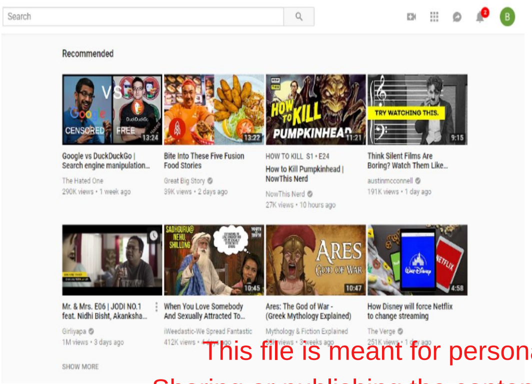
- Help user find item of their interest
- Help item provider deliver their items to right user
- Identify products most relevant to the user
- Personalized content
- Help website improve user engagement.

Some facts about recommendation systems

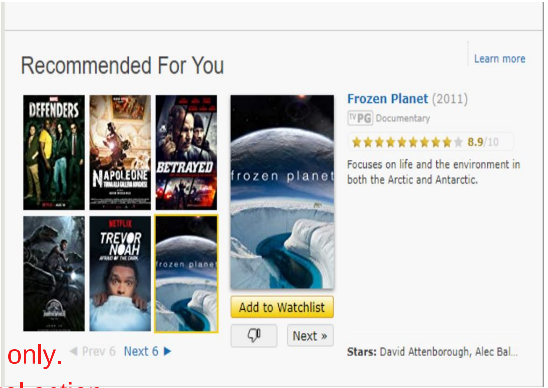
- Netflix – 2/3 rented movies are from recommendation
- Google News – 38% more click-through are due to recommendation
- Amazon – 35% sales are from recommendation

Source: (Celma & Lamere, ISMIR 2007)

Youtube recommendations



IMDB recommendations



Popularity based recommendation systems

It is a type of recommendation systems which suggest products/items based on the popularity or trend. These systems check about the product or movie which are in trend or are most popular among the users and directly recommend those.

Advantages:

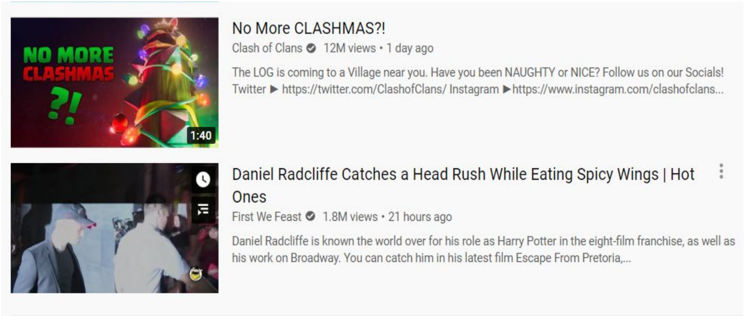
- No cold start problem
- No need for the user's historical data

Disadvantages:

- Not personalized to users
- Only takes popularity in account

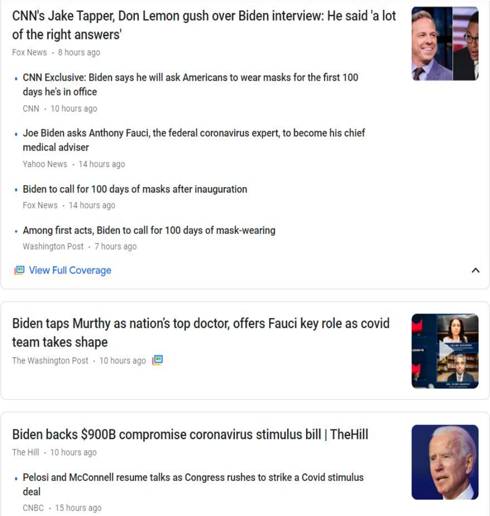
Examples:

1. Google News
2. You tube trending videos



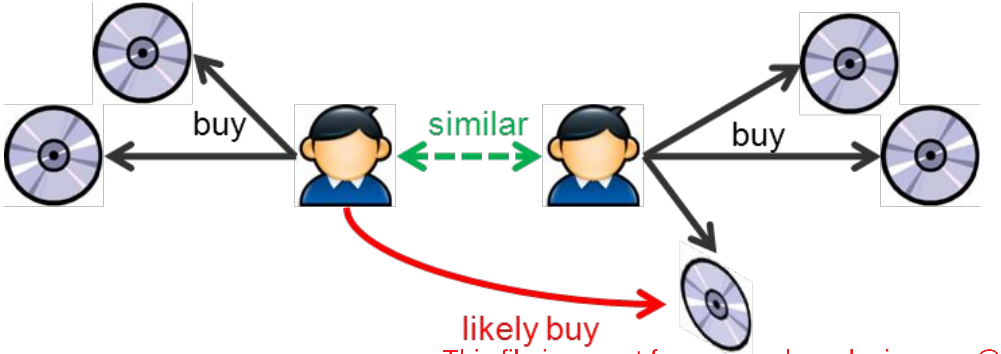
YouTube

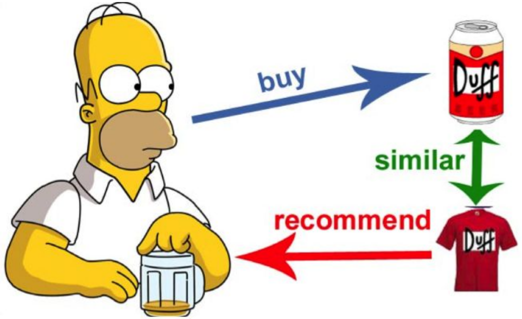
Google News



Collaborative Filtering

- The main idea behind collaborative filtering is that **if a user's likes/dislikes are similar to another user's likes/dislikes, then their tastes are considered similar.** We can use this to recommend a product that other similar users liked.
- It is based on the assumption that if a person who liked something in past will also like it in future.
- It is of two types:
  - **User-User collaborative filtering:** It is based on the search of similar users from the user-item interaction matrix.
  - **Item-Item collaborative filtering:** It is based on the search of similar items from the user-item interaction matrix.





Source: djone

Source: medium

Advantages and Disadvantages of Collaborative Filtering

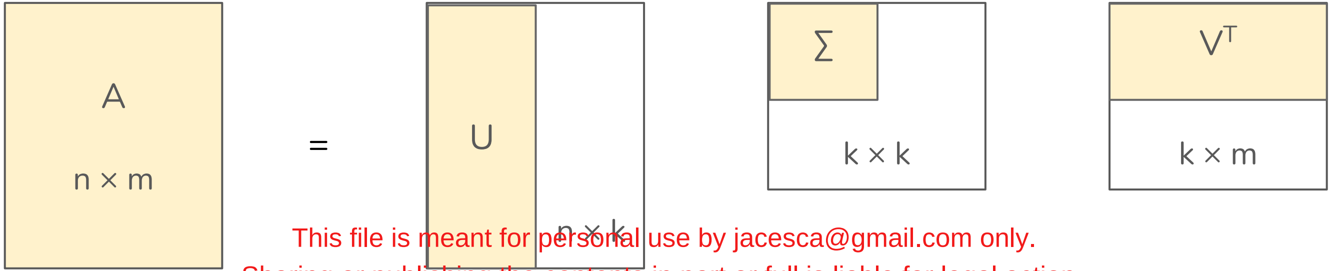
Advantages	Disadvantages
No domain knowledge needed	Cold start problem - can’t handle new data
Personalized to users	Sparsity - Unable to compute ratings if there are a very less number of user preferences
Adaptive to the preferences changes over time	Scalability - Computationally expensive to scale

Matrix Factorization (SVD)

- Sparsity and scalability can be the biggest two challenges with CF methods
- Matrix Factorization decomposes the original sparse matrix to low-dimensional matrices with latent features and less sparsity
- It gives us how much a user is aligned with a set of latent features, and how much a movie fits into this set of latent features
- It uses **Singular Value Decomposition** to factorize the matrix. For a user-movies $n \times m$ matrix, it is given by

$$A = U \Sigma V^T$$

Where, U is an $n \times k$ user-latent feature matrix, V^T is an $k \times m$ movie-latent feature matrix, and Σ is a $k \times k$ diagonal matrix containing the singular values of original matrix, simply representing how important a specific feature is to predict user preference.



Case Study

Happy Learning !