

# Recommender Systems

**Content-based Filtering (CB)**

- Compute similarities between items based on their properties (content)
- Recommend items which are most similar to items a user liked
- Analogous to content-based CF, but similarities are based on item content
- Information about items necessary

## Possible Features:

- Director
- Actor (Top 3)
- Year
- Picture (CNN?)
- Genre
- Runtime
- Language
- Country
- Awards (Oscars won/nominated, Golden Globes won/nominated)
- Tags

## Encoding Techniques

- TFIDF
- One-hot encoding
- Word embeddings

**Collaborative Filtering (CF)**

- use "wisdom of the crowd"
- customers who had similar tastes in the past, will have similar tastes in the future
- no information about items or users necessary

## Model-based CF

**Association rule mining**

- focuses on finding rules that will predict the occurrence of an item based on the occurrences of other items.
- The fact that two items are found to be related means co-occurrence but not causality
- Market basket approach with Apriori algorithm
- "What items do frequently appear together?" (Independent of personale preference profiles)

## Non-parametric approach

**KNN**

- similar to memory-based algorithms, but: we use an unsupervised learning model rather than Pearson correlation or cosine similarity

**Matrix factorization based algorithm**

- Goal: Uncover latent features that explain observed ratings (embeddings)
- Factorization of the user-item ratings matrix

Singular Value Decomposition

Probabilistic Matrix factorization

Non-negative Matrix factorization

## Deep learning

**Multi-layer neural nets**

- extension of matrix factorization method
- matrices don't have to be orthogonal
- input values: user latent features and movie latent features

## Memory-based CF

**User-based CF**

- quantifies the consistencies between the active user and other users
- recommendation of movies that similar users liked
- similar users: Users that rated items similar to the active user
- "Users who are similar to you also liked..."

**Content-based CF**

- quantifies the consistencies between the target item and other items
- recommendation of similar movies to those the user liked based on the ratings of other users
- similar movies: movies that users liked that also liked what the active user liked
- "Users who liked this item also liked..."

**Demographic Filtering**

- Compute similarities between users based on their profiles
- Recommend items which are popular in a user's neighborhood
- Analogous to content-based CF, but similarities are based on item content
- Information about users necessary --> **Not possible**

Cosine similarity or Pearson correlation coefficients  
<https://www.ethanrosenthal.com/2015/11/02/intro-to-collaborative-filtering/>