6.867
Recommender Systems

Fall 2016

The Fall of Search

[the "burden of choice" and
"don't know what you don't know"]

You don't find the content…

…the content finds you!

The long-tail effect

< 20%

NETFLIX

Google

amazon

Alibaba Group
阿里巴巴集团

goodreads
Meet your next favorite book.

YouTube

Spotify

LinkedIn

XING

PANDORA

Microsoft

Yandex

TARGET

criteo labs

tinder

zalando

Quora

OpenTable

$1 million prize

Oct 2006 → Jun 2009

Aim: Predict ratings for movies using training data of (user,movie) pairs of ratings (1-5 stars); improve in-house system by ≥10%!

Training data: 100,480,507 ratings that 480,189 users gave to 17,770 movies

https://en.wikipedia.org/wiki/Netflix_Prize


The ACM Conference Series on
Recommender Systems
HOME
RECSYS 2017
PAST CONFERENCES
HONORS
WIKI
CONTACT
search...
10th ACM Conference on Recommender Systems
RECSYS 2016 (BOSTON)


The ACM Conference Series on
Recommender Systems
HOME
RECSYS 2017
PAST CONFERENCES
HONORS
WIKI
CONTACT
search...
11th ACM Conference on Recommender Systems
RECSYS 2017 (COMO)

Two views on recommendations

Content based

"If you liked this item, you might also like"

Collaborative filtering

"people who liked this item also liked…"
"people similar to you also liked …."

Several other views exist

Learning to rank, choice models, …

"ranked list of preferences (like ml better than xyz)…"

Content based

Features →

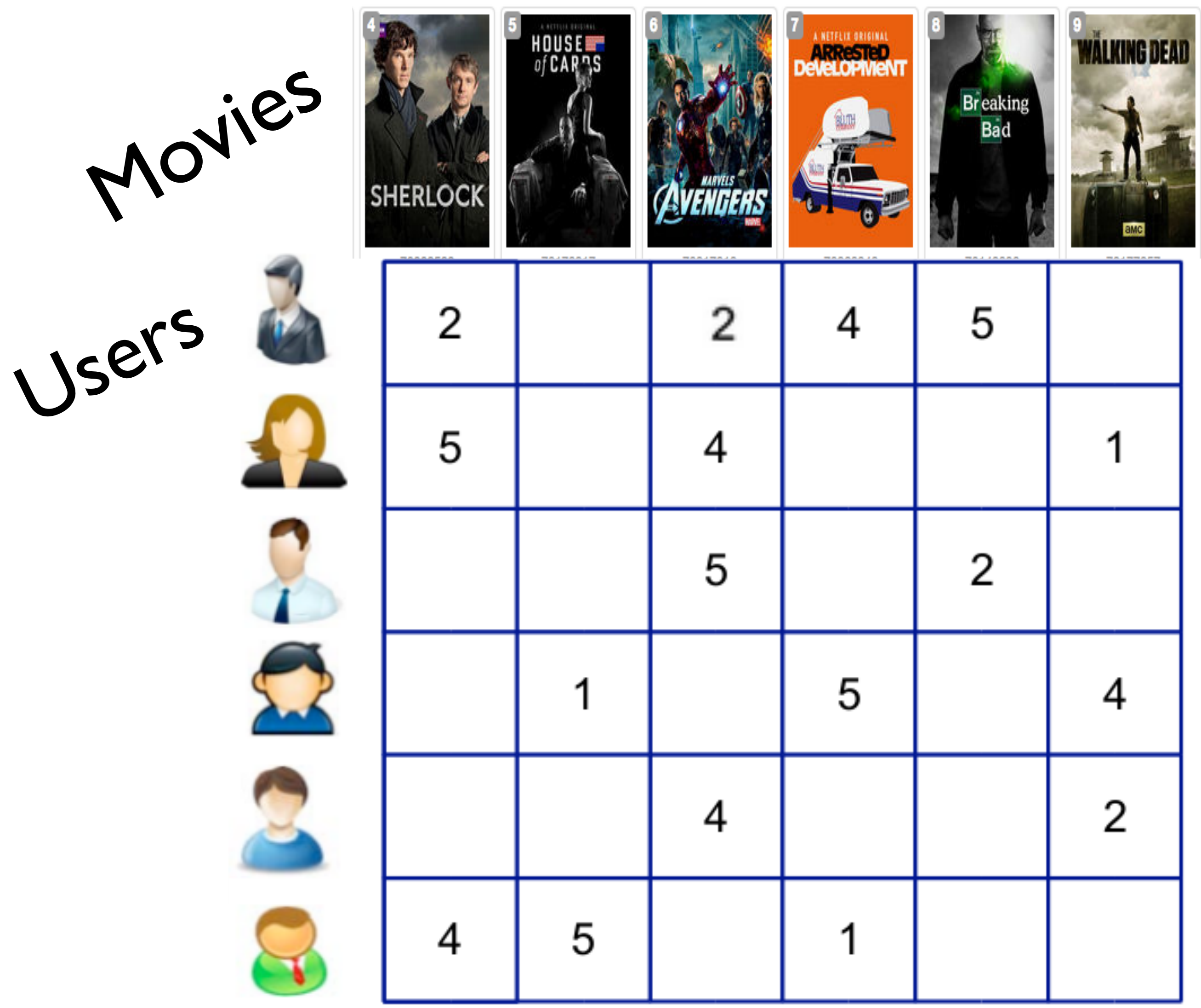
- Genre
- Director
- Actors
- Context
- Astrology

Training data: (Movies, ratings) for a single user

Predict: User's ratings (or like/not-like) for all movies!

More common for text-based products; user models and personalization

Collaborative filtering



2		2	4	5	
5		4			1
		5		2	
	1		5		4
		4			2
4	5		1		

Movies represented by how others have rated them: **no features!**

Various ways of now "filling in the matrix"

Image credit: [X. Amatriain, MLSS'14]

Similarity - sample correlation

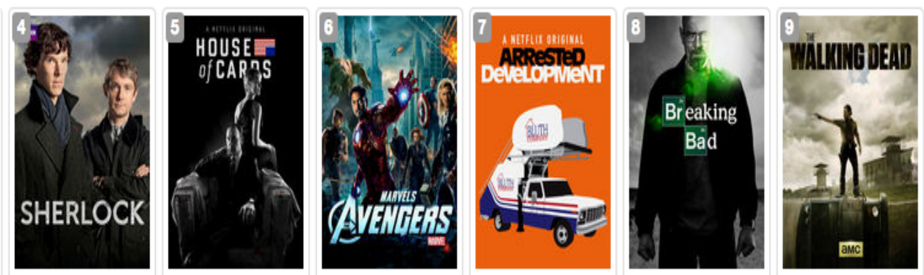
$$\mathrm{sim}(a,b) = \frac{\sum_{j\in M(a,b)}(Y_{aj}-\bar{Y}_a)(Y_{bj}-\bar{Y}_b)}{\sqrt{\sum_{j\in M(a,b)}(Y_{aj}-\bar{Y}_a)^2}\sqrt{\sum_{j\in M(a,b)}(Y_{bj}-\bar{Y}_b)^2}}$$

$$\bar{Y}_a = \frac{1}{|M(a,b)|}\sum_{j\in M(a,b)} Y_{aj}$$

$$\mathrm{sim}(a,b) = \frac{\langle \hat{Y}_a, \hat{Y}_b\rangle}{\|\hat{Y}_a\|\cdot\|\hat{Y}_a\|}$$

$$\hat{Y}_a = [Y_{aj}-\bar{Y}_a]_{j\in M(a,b)}$$

Collaborative filtering via KNN



	2		2	4	5	
b	5		4			1
			5		2	
		1		5		4
a			4			2
	4	5		1		

$$\mathrm{sim}(a,b) = 1$$

Image credit: [X. Amatriain, MLSS'14]

Collaborative filtering via KNN

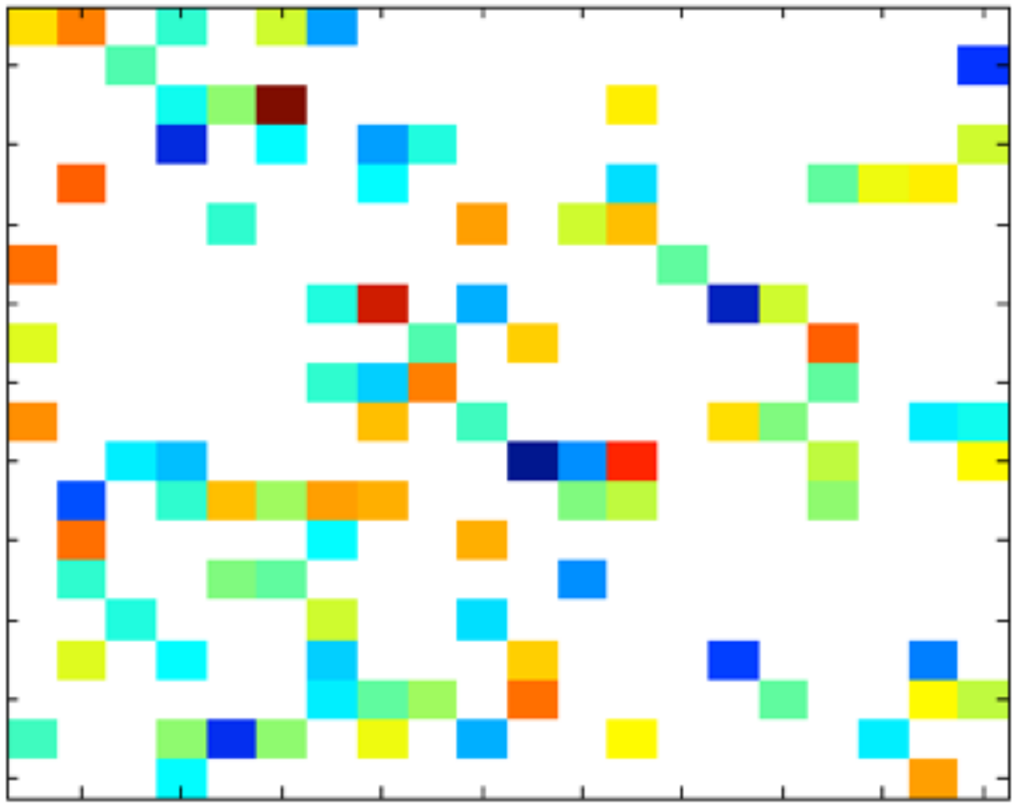
Strengths

- Conceptually simple
- Easy to implement
- Typically few parameters (just K)
- Many improvements possible, eg by designing notion of similarity
- Works well for "stereotypical" users

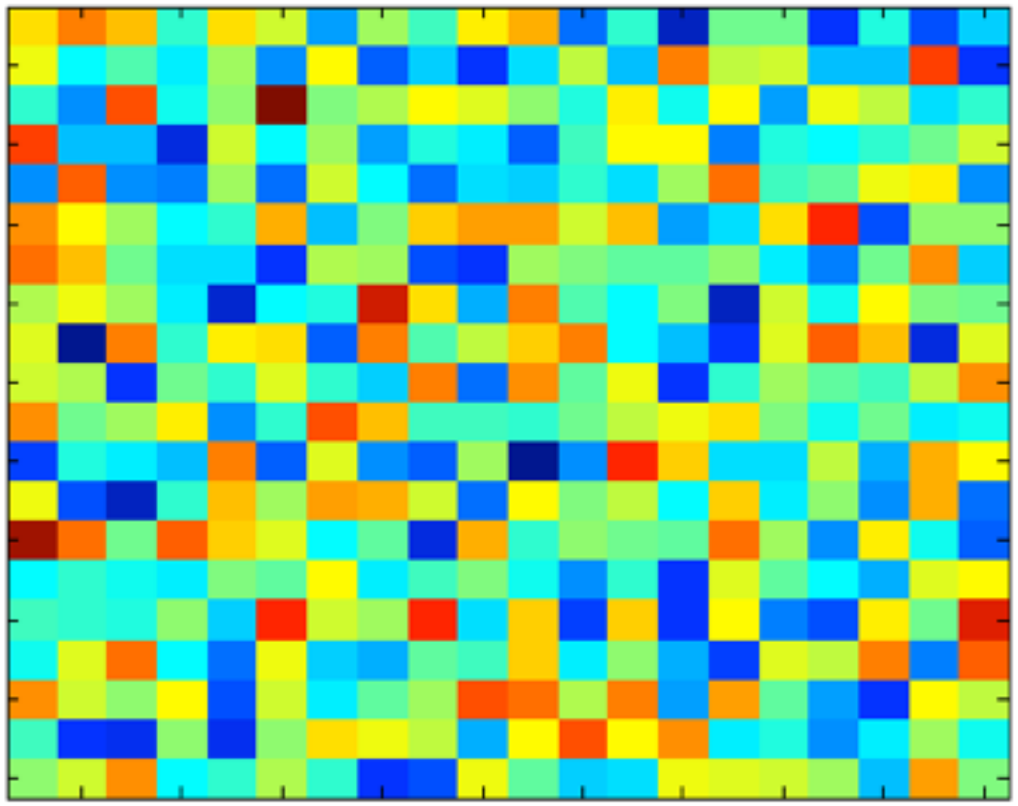
Weaknesses

- Not good for users with mixed tastes (e.g., when user tastes are similar across subsets but diverge / vary for certain other movies)
- Not so good for "diversity"
- Cold-start problem
- Sparsity difficulties
- Scalability of NN

Collaborative filtering: Matrix Factorization



Y



X

A bad idea: trivial regression

2		2	4	5	
5		4			1
		5		2	
	1		5		4
		4			2
4	5		1		

Y

$$\frac{1}{2}\sum_{(a,i)\in M}(Y_{ai}-X_{ai})^2+\frac{\lambda}{2}\sum_{(a,i)\in M}X_{ai}^2$$

$$\hat{X}_{ai}=\begin{cases}\frac{1}{1+\lambda}Y_{ai}, & (a,i)\in M\\ 0 & \text{otherwise}\end{cases}$$

Too many parameters!

Low-rank matrix factorization

$$\min \sum_{(a,i)\in M} (Y_{ai} - X_{ai})^2$$
$$\text{s.t. } \operatorname{rank}(X) \leq k.$$

rank constraint leads to NP-Hard problem
famous "convex relaxation"

$$\operatorname{rank}(X) \leq k \mapsto \left(\|X\|_* := \sum\nolimits_{j=1}^{m} \sigma_j(X)\right) \leq k$$

?

How to solve this problem?

Low-rank MF: AltMin

$$\min_{U,V} \quad F(U,V) := \|P_M(Y) - P_M(UV^T)\|_F^2,$$

$$[P_M(X)]_{ai} = \begin{cases} X_{ai} & (a,i) \in M \\ 0 & \text{otherwise} \end{cases}$$

Observation: Convex in U if V is fixed and vice-versa

Theorem: Under some (academic) assumptions, AltMin initialized with SVD converges to global optimum of this nonconvex problem

Exercise: Compare with an SGD based algorithm!

Other topics, perspectives

Personalization for Google Now (user models)

People Recommendation

Group Recommender Systems

Recommendations within a Social Network

Tensor Factorization

Evaluation of Recommender Systems

Real-time Recommendation of Streamed Data

Interactive Recommender Systems

Links / References

General

http://www.recsyswiki.com/

MLSS 2014: Recommender Systems

see also the "recommended" links on that page ;-)

http://www.recommenderbook.net/

high level overview (a bit dated)

Software

https://github.com/geffy/tffm

factorization machines in TensorFlow

http://www.librec.net

Java library with several algorithms