



the science behind
Netflix recommendations

agenda

1. an introduction to recommendation engines (three types!)
2. how does Netflix make recommendations?
3. how can we code a recommendation engine?

// part 1: an introduction to recommendation engines

types of recommendation engines, and why???

- why do we need recommendation engines + what are some examples?
- three main types of recommendation engines:
 - a. non-personalized
 - b. content-based
 - c. collaborative filtering

non-personalized recommendation engines

The screenshot displays the YouTube homepage with a left-hand navigation menu and a main content area. The navigation menu includes links to Home, Trending, Subscriptions, Library, and History, along with a 'SIGN IN' button and a 'BEST OF YOUTUBE' section with categories like Music, Sports, Gaming, Movies, TV Shows, News, Live, Spotlight, and 360° Video. The main content area features a 'YouTube Music' banner with the text 'Our new music streaming service is here.' and a 'LET'S GO' button. Below this is a 'Trending' section with a grid of video thumbnails, including 'SPIDER-MAN: FAR FROM HOME - Official Trailer', 'HIGHLIGHTS | Canelo Alvarez vs. Daniel Jacobs', 'Ping Pong Trick Shots 5 | Dude Perfect', 'Family Feud Cold Open - SNL', and '\$800 KITCHEN KNIFE'. At the bottom is a 'YouTubeTV' section with the text 'Live TV from 70+ networks' and a 'TRY YOUTUBE TV' button, followed by a row of program thumbnails for 'The Voice', 'GREY'S ANATOMY', and 'SPORTSCENTER'.

YouTube

Search

YouTube Music

Our new music streaming service is here.

LET'S GO

Trending

SPIDER-MAN: FAR FROM HOME - Official Trailer
Sony Pictures Entertainment
31M views • 1 day ago

HIGHLIGHTS | Canelo Alvarez vs. Daniel Jacobs
DAZN USA
3.4M views • 2 days ago

Ping Pong Trick Shots 5 | Dude Perfect
Dude Perfect
8.4M views • 22 hours ago

Family Feud Cold Open - SNL
Saturday Night Live
3M views • 2 days ago

\$800 KITCHEN KNIFE
BuzzFeedVideo
1.5M views • 2 days ago

YouTubeTV Featured

Live TV from 70+ networks

\$49.99 per month
Cancel anytime

TRY YOUTUBE TV

The Voice
Mondays and Tuesdays 8/7c

GREY'S ANATOMY
Thursdays 8/7c

SPORTSCENTER
Tune in Daily 9/8c

content-based recommendation engines

- makes recommendations based on an item's **features**
- based on “similar” items, using other Machine Learning models

movies	Genre	Actor	Director	Year	IMDB	Rotten Tomatoes	...
1							
2							
3							
4							
5							
...							

content-based recommendation engines

- what are some pros and some pitfalls of content-based recommendations?



collaborative filtering -- the utility matrix

- a utility matrix shows user ratings of different items (usually sparse)
- the idea is to **fill in the blanks**, and recommend items with the highest predictions

	Movie 1	Movie 2	Movie ...	Movie N
User 1	1	BLANK	BLANK	3
User 2	BLANK	5	BLANK	3
User 3	BLANK	BLANK	1	BLANK
User 4	2	3	BLANK	BLANK
User 5	BLANK	BLANK	1	BLANK
User 6	4	BLANK	5	BLANK
User 7	BLANK	4	BLANK	BLANK
User ...	BLANK	3	BLANK	BLANK
User m	BLANK	BLANK	BLANK	4



	Movie 1	Movie 2	Movie ...	Movie N
User 1	1	4	2	3
User 2	1	5	3	3
User 3	2.5	2.8	1	3.5
User 4	2	3	2	3.5
User 5	2.5	2.8	1	3.1
User 6	4	1.2	5	1.4
User 7	1	4	2.5	3
User ...	2	3	2	3
User m	1	4	2	4

collaborative filtering

- recommends items **based on ratings of other users**
- different ways to do collaborative filtering:
 - model-based (matrix factorization)
 - modified Singular Value Decomposition

modified SVD for filling in utility matrices

	U1	U2	U3	U4
I1	4	2	3	5
I2	3	2	4	2
I3	?	4	5	4
I4	3	2	4	4

\approx

I1	x	x
I2	x	x
I3	x	x
I4	x	x

latent features

\cdot

U1	U2	U3	U4
x	x	x	x
x	x	x	x

latent features

gradient descent with Alternating Least Squares

I1	1	1
I2	1	1
I3	1	1
I4	1	1

•

U1	U2	U3	U4
1	1	1	1
1	1	1	1

=

	U1	U2	U3	U4
I1	2	2	2	2
I2	2	2	2	2
I3	2	2	2	2
I4	2	2	2	2

OG	U1	U2	U3	U4
I1	4	2	3	5
I2	3	2	4	2
I3	?	4	5	4
I4	3	2	4	4

RMSE = 1.75

gradient descent with Alternating Least Squares

I1	x	1
I2	1	1
I3	1	1
I4	1	1

•

U1	U2	U3	U4
1	1	1	1
1	1	1	1

=

	U1	U2	U3	U4
I1	x+1	x+1	x+1	x+1
I2	2	2	2	2
I3	2	2	2	2
I4	2	2	2	2

OG	U1	U2	U3	U4
I1	4	2	3	5
I2	3	2	4	2
I3	?	4	5	4
I4	3	2	4	4

Finding x to minimize:

$$(4-(x+1))^2 + (2-(x+1))^2 + (3-(x+1))^2 + (5-(x+1))^2$$

setting $d/dx = 0$, $x = 2.5$

gradient descent with Alternating Least Squares

I1	2.5	1
I2	1	1
I3	1	1
I4	1	1

•

U1	U2	U3	U4
1	1	1	1
1	1	1	1

=

	U1	U2	U3	U4
I1	3.5	3.5	3.5	3.5
I2	2	2	2	2
I3	2	2	2	2
I4	2	2	2	2

OG	U1	U2	U3	U4
I1	4	2	3	5
I2	3	2	4	2
I3	?	4	5	4
I4	3	2	4	4

RMSE = 1.58!!

collaborative filtering -- pros and cons

- personalized for each user
- computationally heavy
- popularity bias
- the **cold start** problem



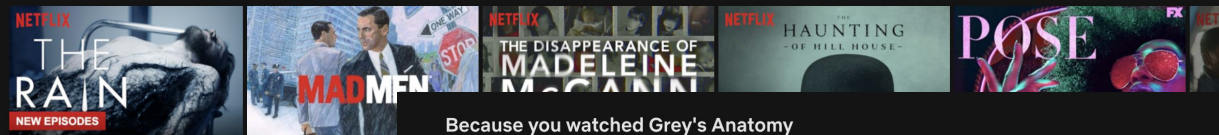
part 2: how does Netflix
make recommendations?

the Netflix algorithm

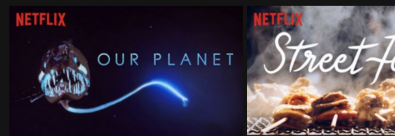
- Netflix uses a **hybrid** of content-based and collaborative filtering
- content based: tagging
- collaborative filtering: user patterns, user similarities

the Netflix algorithm

Emotional TV Shows



Documentaries



Because you watched Grey's Anatomy



Suspenseful TV Shows



Crime TV Shows



Because you watched Russian Doll



// part 3: coding our own
recommendation engine!

stuff we've learned

recommendation engines!!!

1. non-personalized
2. content-based
3. model-based collaborative filtering with SVD

code-along: <https://github.com/yishuen/meetup-movie-recommender>