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

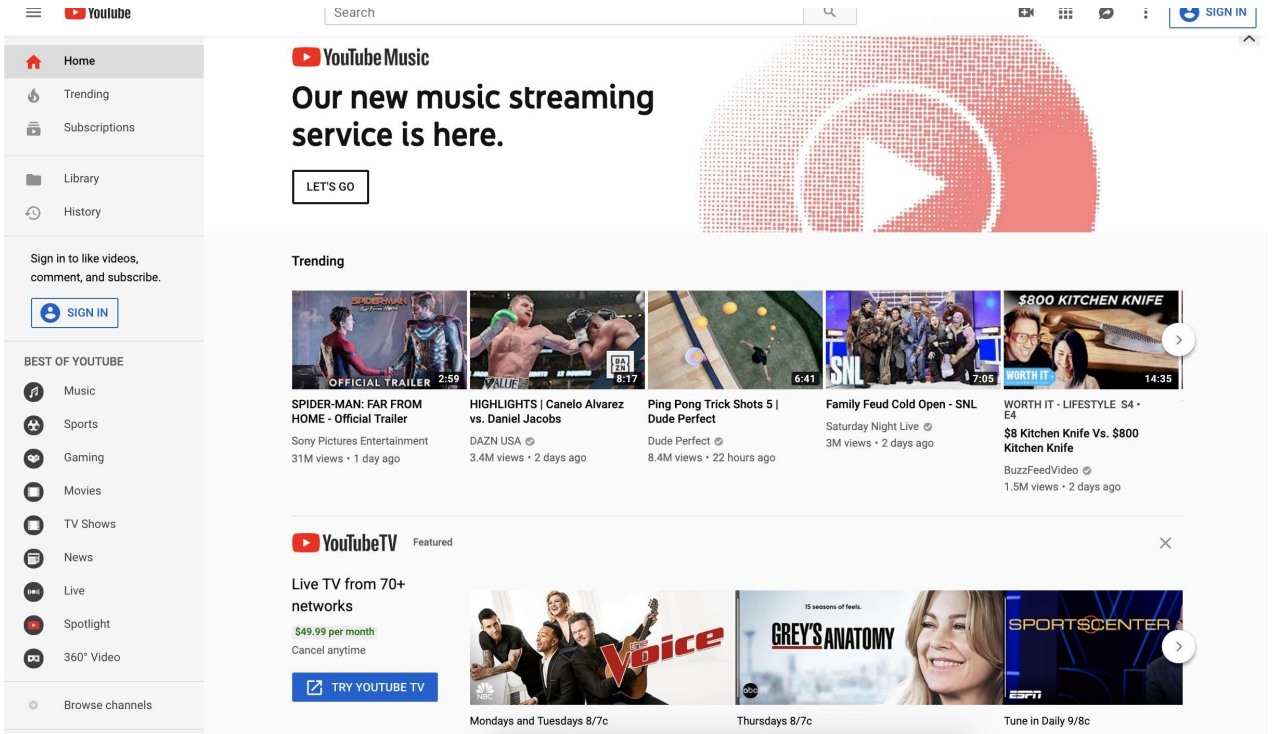
types of recommendation engines, and **why**???

- 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



content-based recommendation engines

- makes recommendations based on an item's **features**

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

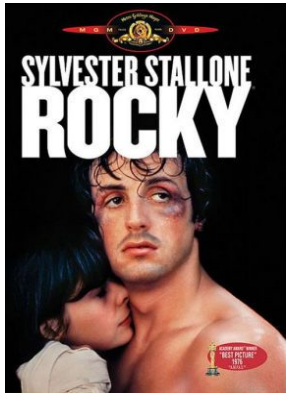
content-based recommendation engines

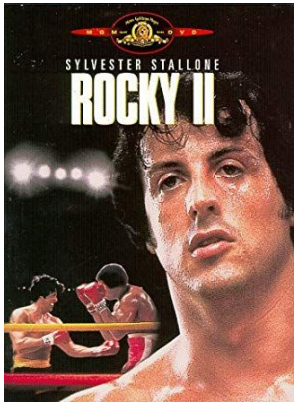
- makes recommendations based on an item's **features**

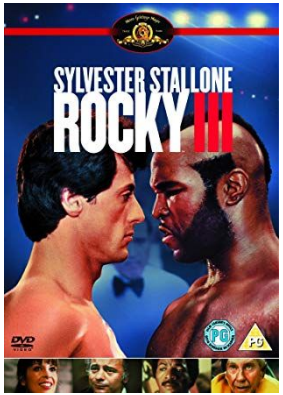
movies	Genre	Actor	Director	Year	IMDB	Rotten Tomatoes	...
1 Rocky	Drama	Stallone	Avildsen	1976	8.1	94%	
2 Rocky II	Drama	Stallone	Stallone	1979	7.2	77%	
3 First Blood (Rambo)	Action	Stallone	Kotcheff	1982	7.7	87%	
4 The Expendables	Action	Stallone	Stallone	2010	6.5	42%	
5 Judge Dredd	Action	Stallone	Cannon	1995	5.6	17%	
...							

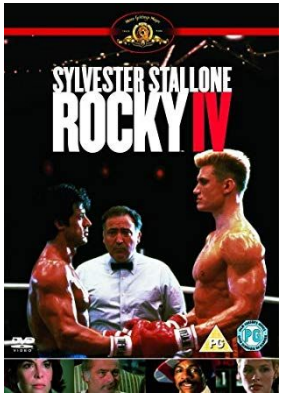
content-based recommendation engines

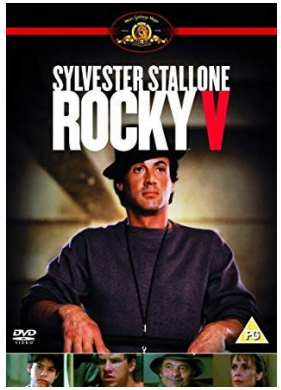
- what are some pros and some pitfalls of content-based recommendations?











collaborative filtering

- recommends items **based on user feedback**

- different ways to do collaborative filtering:
  - Singular Value Decomposition (SVD)
  - Latent Dirichlet Allocation (LDA)
  - Bayesian Networks
  - Cosine similarity

implicit vs explicit feedback

- Collaborative filtering systems require feedback!

Explicit

- Given directly by user - (e.g. star rating, upvotes)
- Need to ask for user input
- Expensive / more difficult to collect

Implicit

- Given indirectly (e.g. movie user watched, demographic data)
- Privacy concerns in collection
- Large quantities of implicit data
- Can be difficult to interpret

collaborative filtering recommendation engines

- make recommendations based on **implicit feedback**

user	Alien	Predator	Starship Troopers	Fight Club	Inception	Memento	...
1 John	1	1	0	0	1	0	
2 Alyssa	1	1	0	1	0	0	
3 Nick	0	0	1	0	0	1	
4 Jill	1	0	1	0	0	1	
5 Melissa	0	0	0	1	1	1	
...							

collaborative filtering recommendation engines

- make recommendations based on **explicit feedback**

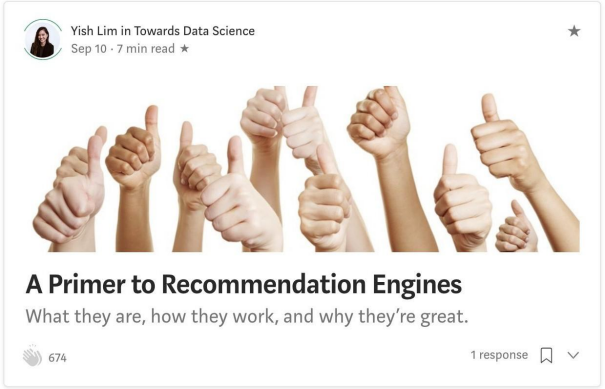
user	Alien	Predator	Starship Troopers	Fight Club	Inception	Memento	...
1 John	5	4	NA	NA	2	NA	
2 Alyssa	5	4	NA	5	NA	NA	
3 Nick	NA	NA	5	NA	NA	4	
4 Jill	1	NA	4	NA	NA	4	
5 Melissa	NA	NA	NA	5	1	5	
...							

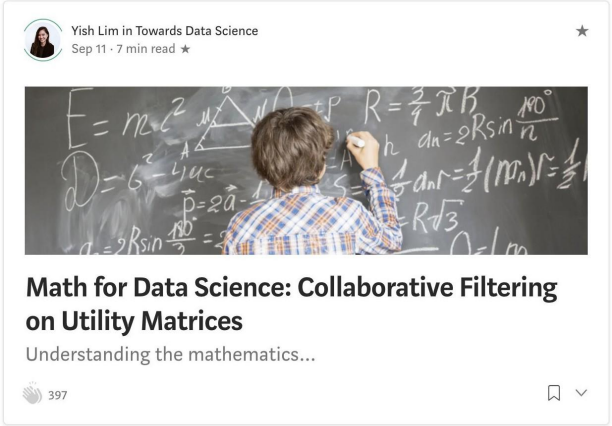
collaborative filtering -- pros and cons

- personalized for each user!

- computationally heavy
- popularity bias

- **grey sheep** problem
- the **cold start** problem



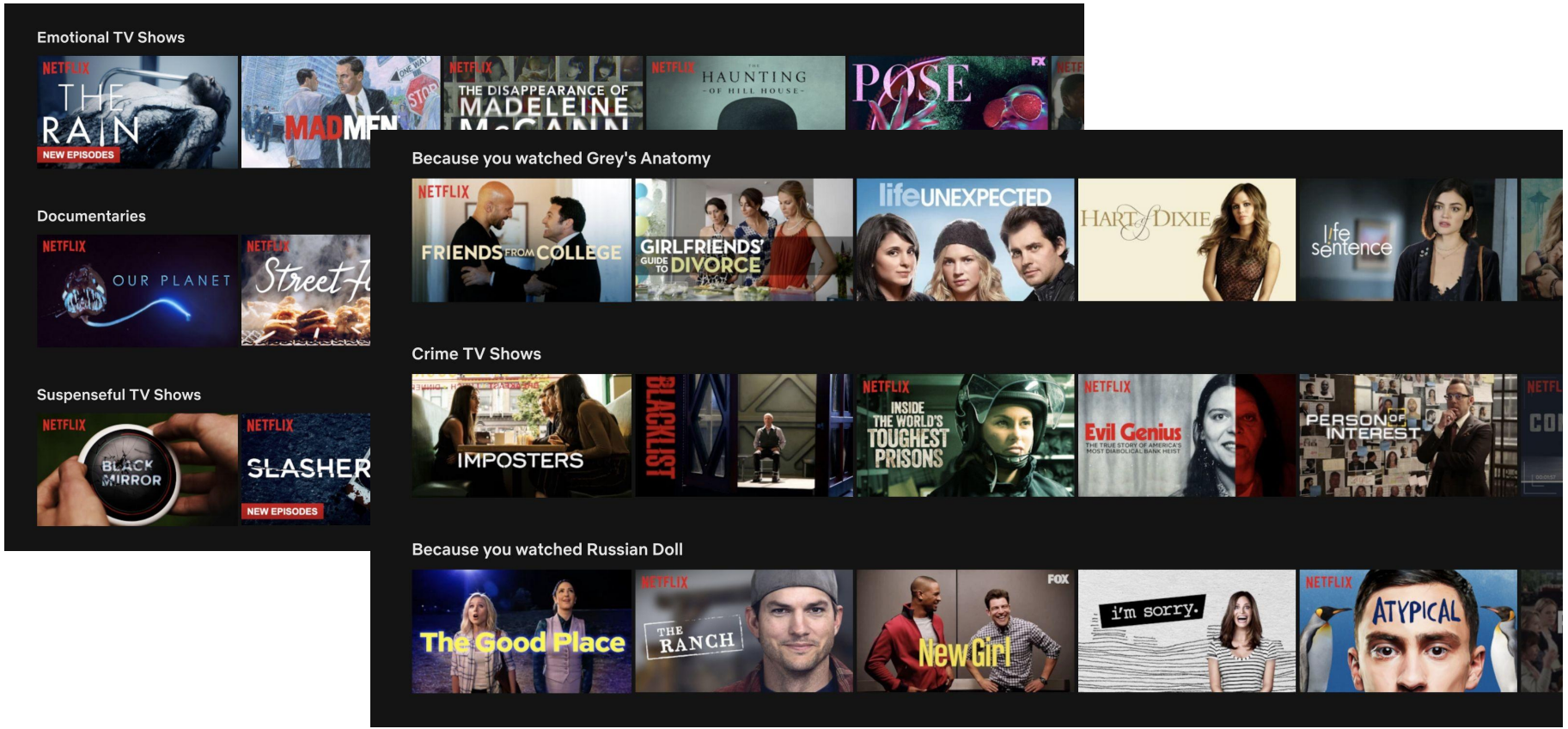


part 2: how does Netflix make recommendations?

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



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