

Unsupervised Learning

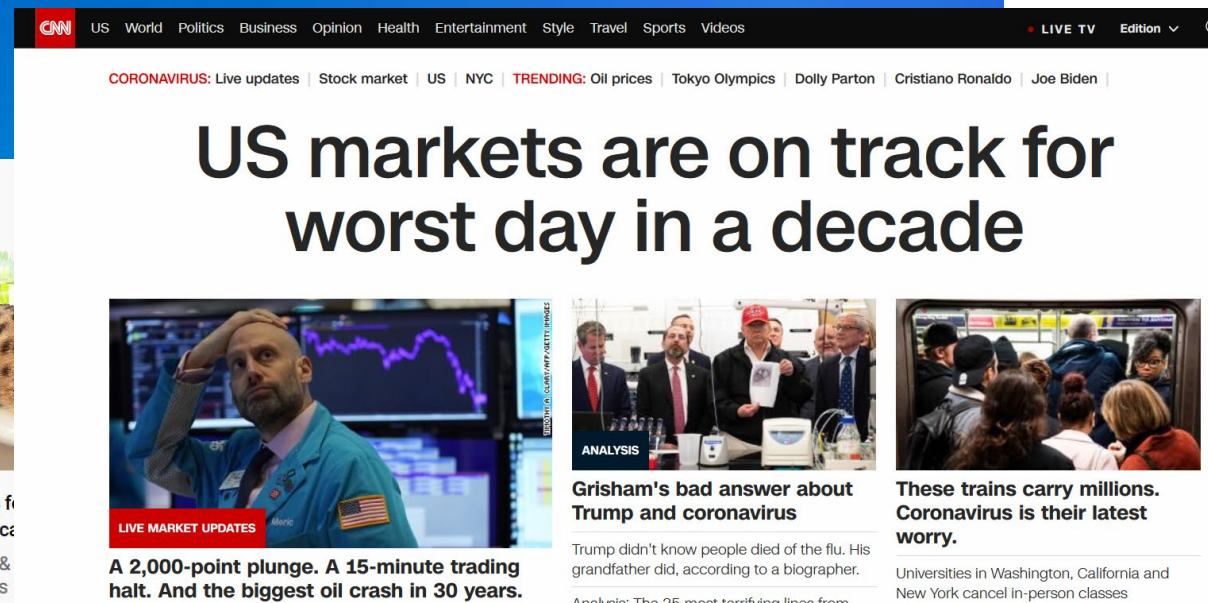
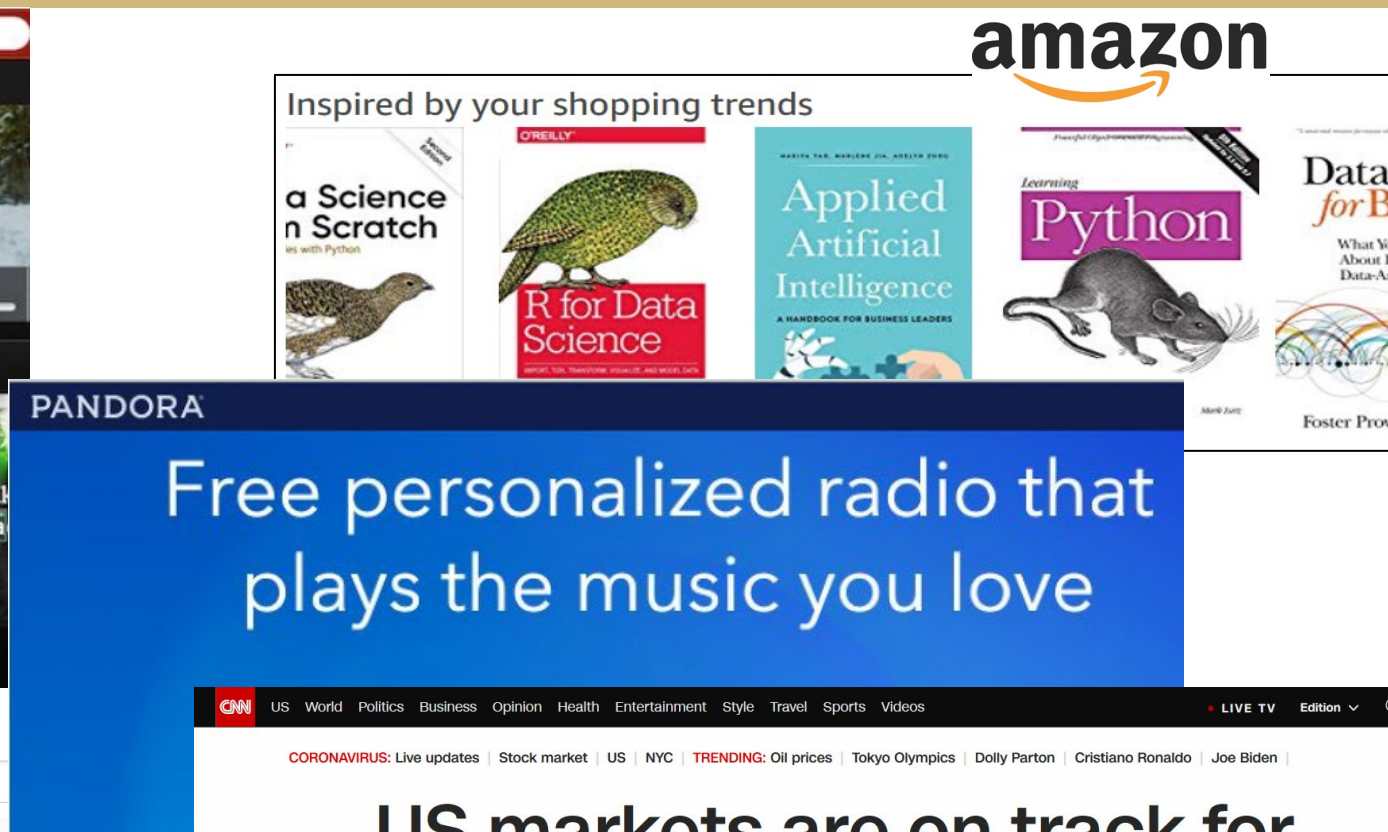
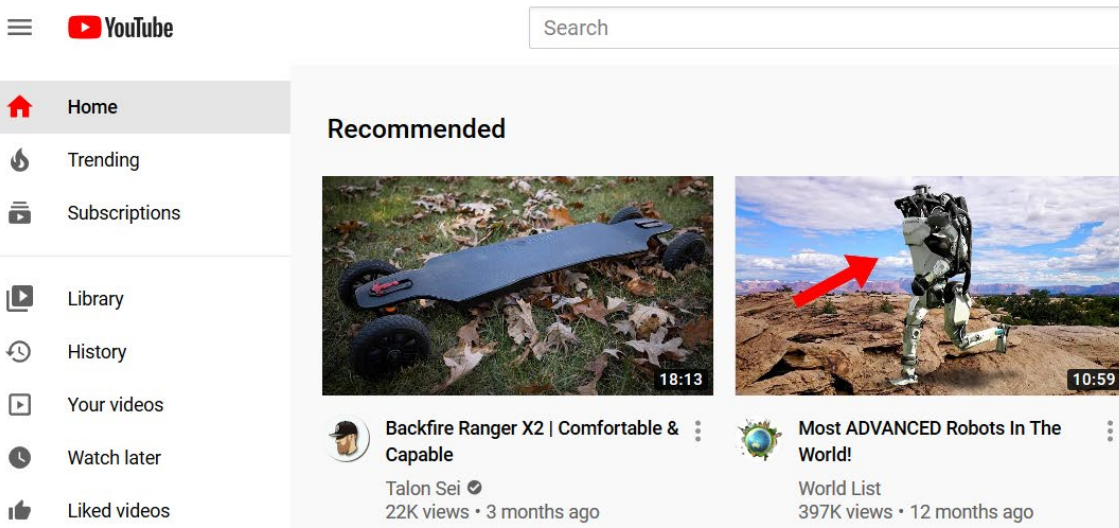
Geena Kim



Recommender System

- What is Recommender System
- Content-based Filtering
- Collaborative Filtering
- Matrix Factorization

What is a Recommender System?



Business Goals

What will the user **Like**?

What will the user **Buy**?

What will the user **Click**?

Goal of recommender

- Predict missing ratings
- May be sufficient to just predict a subset of items with high expected rankings
- May be sufficient to just predict general trends, such as *trending* news

Recommendation approaches

Popularity:

- Make the **same** recommendation to **every** user, based only on the popularity of an item.
- E.g. Twitter “Moments”

Content-based (Content filtering):

- Predictions are made based on the properties/characteristics of an item.
- User behavior is **not** considered. (There can be hybrid)
- E.g. Pandora Radio

Recommendation approaches

Collaborative filtering:

- Only consider past user behavior.
(**not** content properties...)
- User-User similarity
- Item-Item similarity
- E.g.
 - Netflix & Amazon Recommendations,
 - Google Ads,
 - Facebook Ads, Search, Friends Rec., News feed, Trending news, Rank Notifications, Rank Comments

Similarity-based Collaborative Filtering:

- Use similarity metric between users or items

Matrix Factorization Methods:

- Find latent features/factors

Content-Based Filtering

- Creates profile of each user and items
- Need to collect user demographics or questionnaire
- Need domain-specific info about the items
- Features are hand-engineered by the domain experts

Content-Based Filtering Example

pandora® | Music Genome Project

About

Contact

Press

Management

Board

About The Music Genome Project®

The Music Genome Project powers Pandora. It's the most comprehensive analysis of music ever undertaken.

For over a decade, we've been gathering musical knowledge to bring you the best, most personalized listening experience out there.

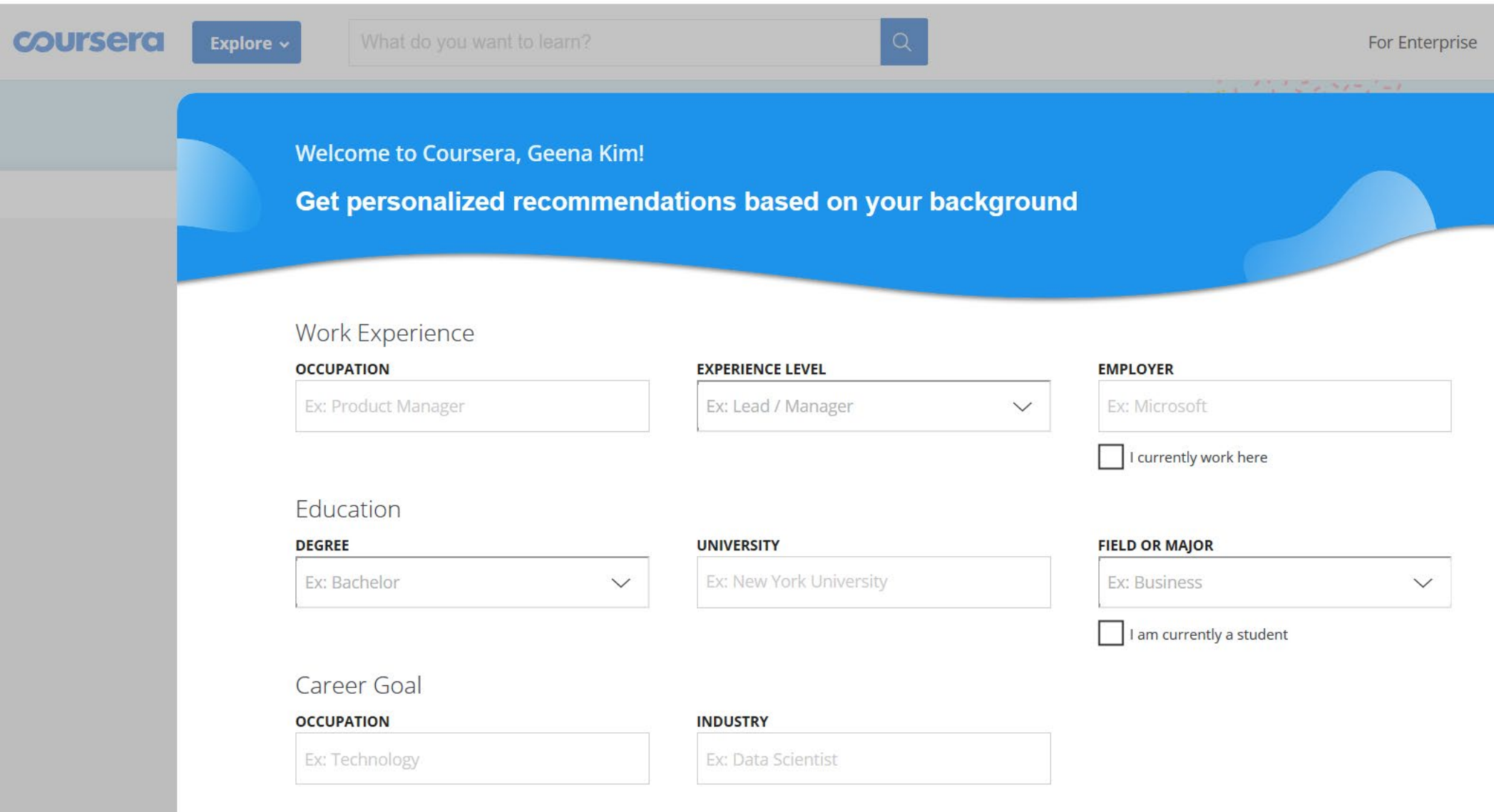
We believe each individual has a unique relationship with music – no one has tastes that are exactly the same. So delivering a great experience to every listener requires a broad and deep understanding of music.

Our team of trained musicologists has been listening to music across all genres and decades, including emerging artists and new releases, studying and collecting musical details on every track– 450 musical attributes altogether.

The result of all our work is a personalized listening experience filled with both old favorites and new discoveries.

User profiling

Oftentimes Recommender systems collect user data



The image shows the Coursera website's user profile page. At the top, the Coursera logo is on the left, followed by an 'Explore' button with a dropdown arrow. In the center is a search bar with the placeholder text 'What do you want to learn?' and a magnifying glass icon. On the right is a link for 'For Enterprise'. Below the header is a blue banner with the text 'Welcome to Coursera, Geena Kim!' and 'Get personalized recommendations based on your background'. The main content area contains three sections: 'Work Experience', 'Education', and 'Career Goal'. Each section has several input fields for user data.

Work Experience

OCCUPATION
Ex: Product Manager

EXPERIENCE LEVEL
Ex: Lead / Manager

EMPLOYER
Ex: Microsoft

☐ I currently work here

Education

DEGREE
Ex: Bachelor

UNIVERSITY
Ex: New York University

FIELD OR MAJOR
Ex: Business

☐ I am currently a student

Career Goal

OCCUPATION
Ex: Technology

INDUSTRY
Ex: Data Scientist

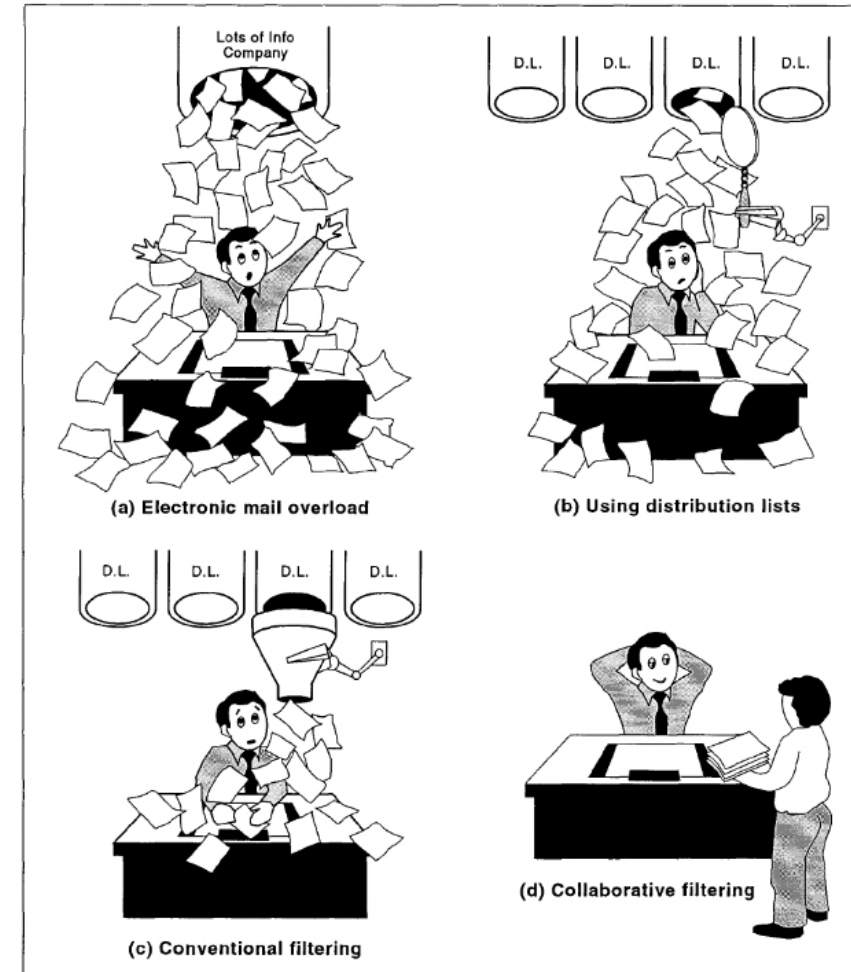
Collaborative Filtering

- No need of hand-engineered features
- Domain-free
- Can learn latent info that are hard to profile using content filtering
- Generally more accurate about user preference
- May suffer from “cold-start” problem

(

Tapstry

The origin of the name: filtering documents in emails



Goldberg et. al., (1992)

Collaborative Filtering Approaches

Memory-based

- Customers who bought this item also bought

Using Similarity

- item-item similarity
- user-user similarity
- Can be combined with clustering

Using Latent factor modeling

- Matrix Factorization
- Tries to explain the rating patterns

Other

- Deep learning

Similarity

What does the data look like?

	HP1	HP2	HP3	TW	SW1	SW2	SW3	
<i>User</i>	<i>A</i>	4	.	.	5	1	?	
	<i>B</i>	5	5	4	.	?	.	
	<i>C</i>	?	.	?	2	4	5	
	<i>D</i>	?	?	3	?	?	3	

Utility Matrix

User explicitly rate the movie or product

Explicit ratings

Similarity

What does the data look like?

User	Item				
	A	B	C	D	...
	Al	0	1	0	1
	Bob	0	0	1	0
	Cat	0	1	1	1
	Dan	1	0	0	1
	Ed	0	1	0	0
	...				

User buy or not buy the product

Implicit ratings

Similarity

Handwritten notes on the table:

- Red arrow pointing to the first column: $100M$
- Red arrow pointing to the first row: m
- Red arrow pointing to the first cell: n
- Red arrow pointing to the first cell: $600M$

User	Item				
	A	B	C	D	...
	Al	1	?	2	?
	Bob	?	2	3	4
	Cat	3	?	1	5
	Dan	?	2	?	?
	Ed	2	?	?	1
	...				

$m(2 \sim 3(m^2))$
~~mn?~~

User-user similarity: $O(m^2 \times n)$

Item-item similarity: $O(n^2 \times m)$

Calculation cost

Let:
 $m = \text{\#users,}$
 $n = \text{\#items}$

User-User: $O(m^2 n)$
Item-Item: $O(m n^2)$

Similarity

$0 \sim 1$

Distance-based

- Manhattan distance $c=1$
- Euclidean distance $c=2$
- Minkowski distance

$0 \sim \text{inf}$

$$\text{similarity}(a, b) = \frac{1}{1 + \text{dist}(a, b)}$$

$\left(\sum_i (x_A^i - x_B^i)^c \right)^{1/c}$

\downarrow
~~120~~

1

0

Similarity

Pearson Correlation

$$\frac{\text{cov}(a, b)}{\text{std}(a) * \text{std}(b)} = \frac{\sum_i (a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_i (a_i - \bar{a})^2} \sqrt{\sum_i (b_i - \bar{b})^2}}$$

$$\text{similarity}(a, b) = 0.5 + 0.5 * \text{pearson}(a, b)$$

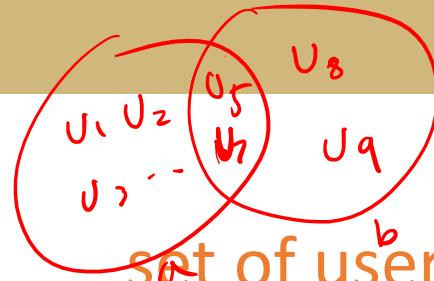
Similarity

✓ Cosine similarity

$$\frac{a \cdot b}{||a|| ||b||} = \frac{\sum_i a_i b_i}{\sqrt{\sum_i a_i^2} \sqrt{\sum_i b_i^2}} \quad -1 \sim 1$$

$$\text{similarity}(a, b) = 0.5 + 0.5 * \cos(\theta_{a,b})$$

Similarity



set of users who rated item a

Jaccard similarity

$$\text{similarity}(a, b) = \frac{|U_a \cap U_b|}{|U_a \cup U_b|}$$

Similarity Matrix

	item 1	item 2	item 3	...
item 1	1	0.3	0.2	...
item 2	0.3	1	0.7	...
item 3	0.2	0.7	1	...
...

similarity measures



Similarity Matrix

User	Item				
	A	B	C	D	...
	Al	1	?	2	?
	Bob	?	2	3	4
	Cat	3	?	1	5
	Dan	?	2	?	?
	Ed	2	?	?	1
	...				

	item 1	item 2	item 3	...
item 1	1	0.3	0.2	...
item 2	0.3	1	0.7	...
item 3	0.2	0.7	1	...
...

$$\text{rating}(u, i) = \frac{\sum_{j \in I_u} \text{similarity}(i, j) * r_{u,j}}{\sum_{j \in I_u} \text{similarity}(i, j)}$$

I_u = set of items rated by user u

$r_{u,j}$ = user u 's rating of item j

Latent factor model

User	Item				
	A	B	C	D	...
	Al	1	?	2	?
	Bob	?	2	3	4
	Cat	3	?	1	5
	Dan	?	2	?	?
	Ed	2	?	?	1
	...				

$$R_{m \times n} \approx U_{m \times k} V_{k \times n}$$
$$r_{ij} \approx u_{i:} \cdot v_{:j}$$

k:
of latent features
is a hyperparameter

UV decomposition

$$\begin{bmatrix} 5 & 2 & 4 & 4 & 3 \\ 3 & 1 & 2 & 4 & 1 \\ 2 & & 3 & 1 & 4 \\ 2 & 5 & 4 & 3 & 5 \\ 4 & 4 & 5 & 4 & \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \\ u_{31} & u_{32} \\ u_{41} & u_{42} \\ u_{51} & u_{52} \end{bmatrix} \times \begin{bmatrix} v_{11} & v_{12} & v_{13} & v_{14} & v_{15} \\ v_{21} & v_{22} & v_{23} & v_{24} & v_{25} \end{bmatrix}$$

References

- [1] Goldberg, D., Nichols, D., Oki, B., & Terry, D. (1992). Using collaborative filtering to weave an information tapestry. *Communications of the Association of Computing Machinery*, 35(12), 61–70.
- [2] Leskovec, J., Rajaraman, A., & Ullman, J. D., Mining Massive Datasets Ch.9 Recommender Systems
<http://infolab.stanford.edu/~ullman/mmds/ch9.pdf>