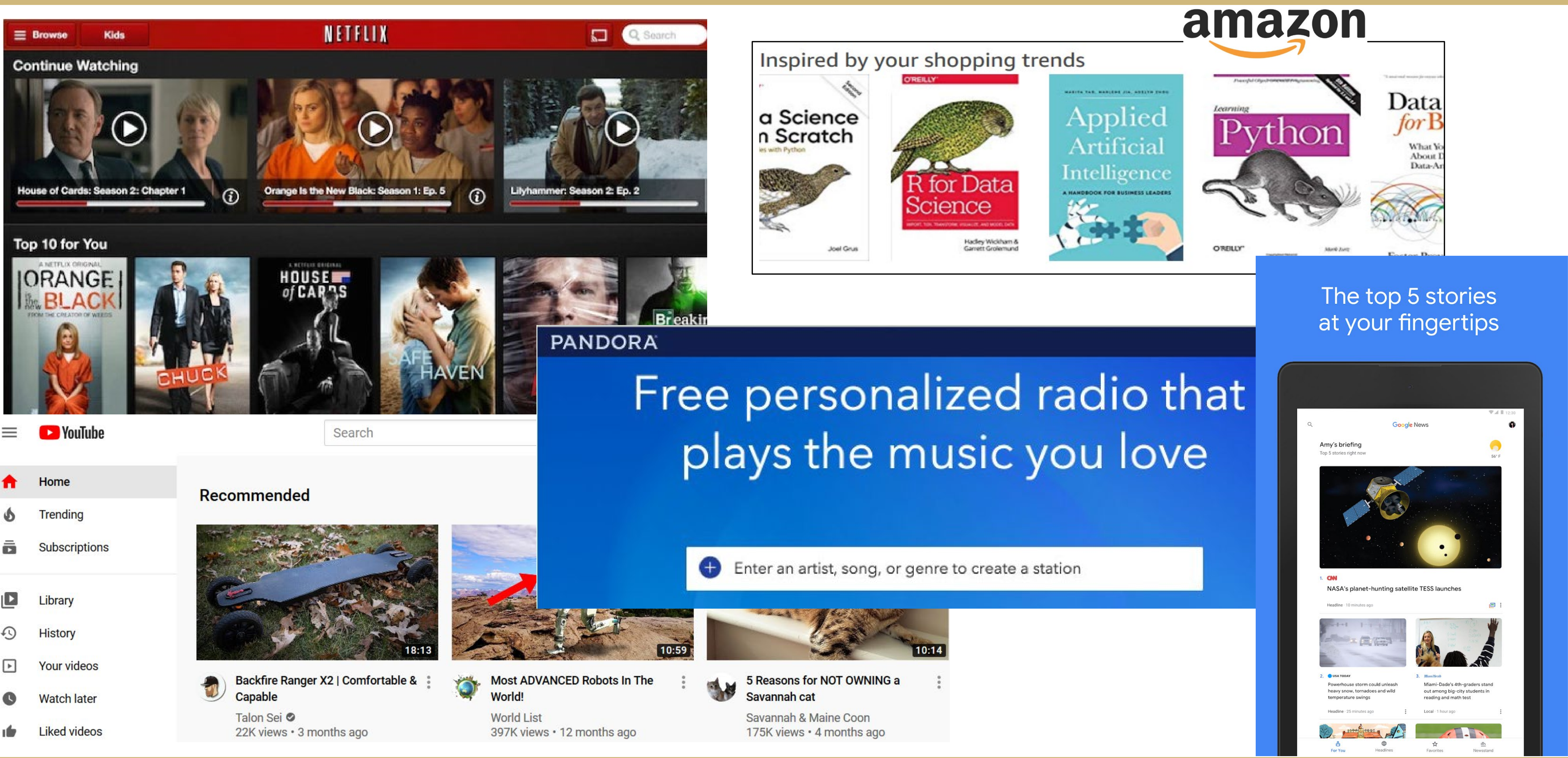
Recommender Systems

Geena Kim

What is a recommender system?



Recommendation approaches

Popularity
- Recommend the most popular or trending item(s) to everyone.

Content-based
- Items are similar if their attributes are similar
- Often hand-engineered (domain-specific) attributes

Collaborative filtering
- Recommends items chosen by similar users
- domain-free

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.

Item features examples

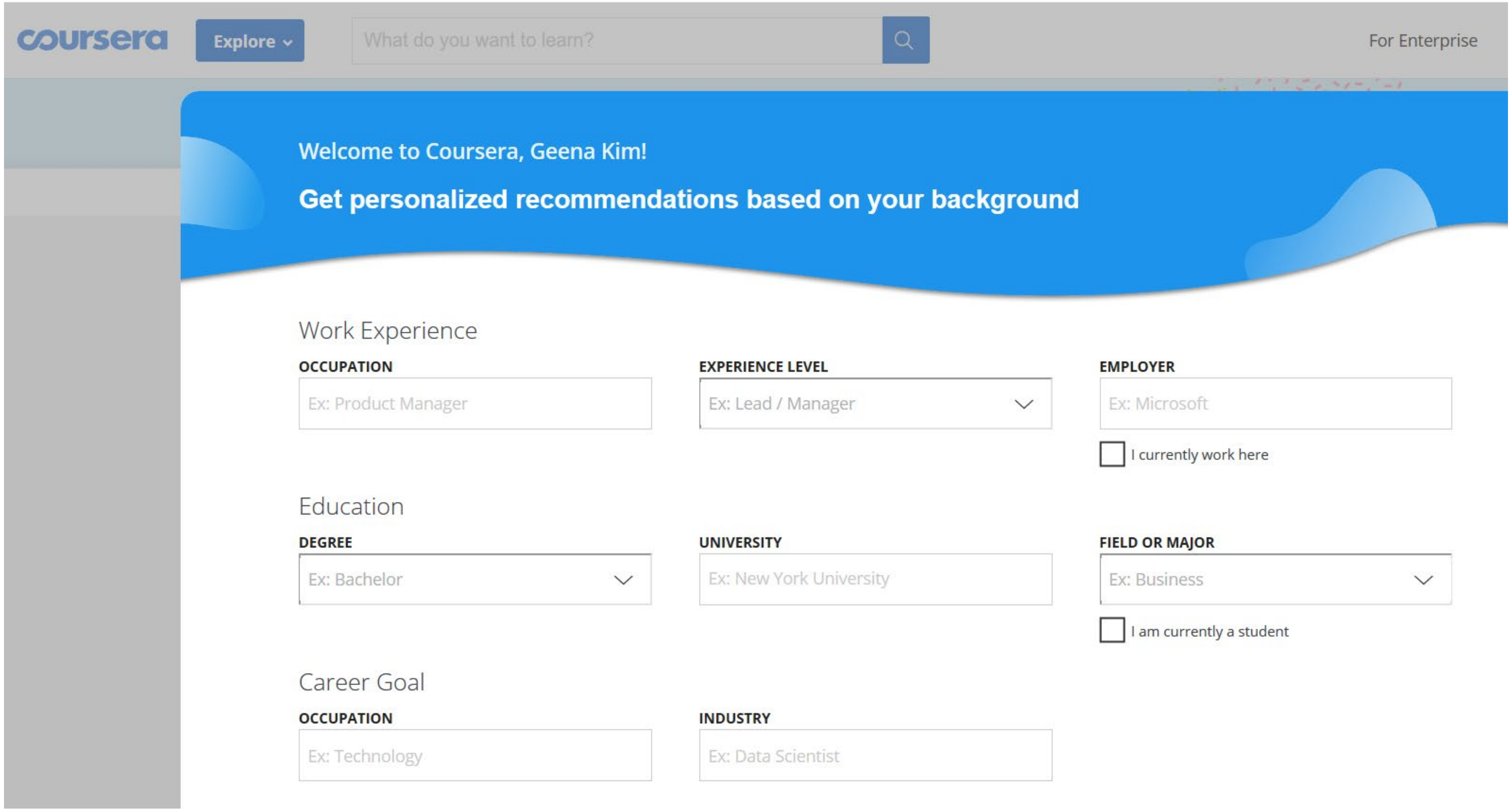
Laptop features: Hard disk size, CPU speed, RAM size, Weight, monitor size, GPU specs

Shirt features: Size, color, fabric, style, collar, finish, sleeve

Movie features: actors, director, year, genre, series, average rating

User profiling

Oftentimes Recommender systems collect user data



Collaborative Filtering

- No need of hand-engineered features
- Domain-free
- Learns from also other users’ interaction with items
- May suffer from **cold-start** problem

Collaborative Filtering Approaches

Memory-based

- Customers who bought this item also bought ….

Using Similarity

- item-item similarity
- user-user similarity

Using Latent factor modeling

- Matrix Factorization

Other

- Supervised approaches, graphs

Utility Matrix

What does the data look like?

	Star Wars I	Star Wars II	Squid Game	Lord of …	Harry Potter I
Amy			2		5
Bob		3	5		
Cathy	1			4	
Dave	4	5		5	

Utility Matrix

User explicitly rate products

Explicit ratings

What does the data look like?

	Item 1	Item 2	Item 3	Item 4	Item 5
A			1		1
B		1	1		
C	1			1	
D	1	1			1

User buy or not buy the product

Implicit ratings

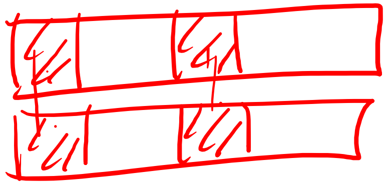
	Item 1	Item 2	Item 3	Item 4	Item 5
A			1		1
B		1	1		1
C	1			1	
D		1	1		1

User-user similarity

Item-item similarity

Similarity measures

Cosine similarity



$$\|\vec{a}\| = \sqrt{\sum_i a_i^2}$$

$$\cos(a,b) = \frac{a \cdot b}{||a||\,||b||}$$

$-1 \sim 1$

$0 \sim 1$

$$\text{sim}(a,b) = 0.5 + 0.5 \times \cos(\text{a,b})$$

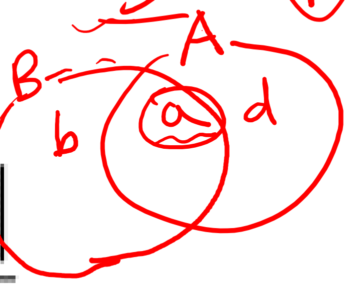
Similarity measures

Jaccard similarity

$$\begin{array}{c|cccc} & a & b & c & d \\ \hline A & 1 & 0 & 0 & 1 \\ B & 1 & 1 & 0 & 0 \end{array}$$

set of users who rated item a

$$\text{Jaccard}(a, b) = \frac{|S_a \cap S_b|}{|S_a \cup S_b|}$$



a. b. d

$= \frac{1}{3}$

$\rightarrow$ 0 ~ 1

Similarity measures

Distance-based

- Manhattan distance
- Euclidean distance
- Minkowski distance

$\sum_i |x_i^a - x_i^b|$

$\sqrt{\sum_i (x_i^a - x_i^b)^2}$

$\left(\sum_i (x_i^a - x_i^b)^N \right)^{1/N}$

$0 \sim \infty$

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

$0 \sim 1$

Similarity measures

Pearson Correlation

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

$-1 \sim 1$

$$\mathrm{sim}(a,b) = 0.5 + 0.5 \times \mathrm{Pearson}(a,b)$$

$0 \sim 1$

Examples- Cosine

	M1	M2	M3	M4
A	5.0	~~NaN~~ 3	1.0	4.0
B	2.0	3.0	5.0	~~NaN~~ 3
C	4.0	4.0	~~NaN~~ 3	4.0

−3

A	+2	0	−2	+1
B	−1	0	+2	0
C	+1	+1	0	+1

$$\cos(A,B) = \frac{5\times 2 + \cancel{0\times 3} + 1\times 5 + \cancel{4\times 0}}{\sqrt{5^2+0^2+1^2+4^2}\times\sqrt{2^2+3^2+5^2+0^2}}$$

$= 0.375\ldots$

$\cos(A,C) = \sim 0.8$

$\cos(B,C) = \sim 0.47$

$$\cos(A,B) \; \frac{5\times 2 + 3\times 3 + 1\times 5 + 4\times 3}{\sqrt{5^2+3^2+1^2+4^2}\sqrt{2^2+3^2+5^2+3^2}} = 0.24$$

$$\cos(A,B) = \frac{2\times -1 + 0 - 2\times 2 + 0}{\sqrt{2^2+0+-2^2+1^2}\sqrt{-1^2+0+2^2+0}} = \sim -0.89$$

$\cos(A,C) \sim 0.58$

$\cos(B,C) \sim -0.26$

Examples- Cosine

	M1	M2	M3	M4	M5	
A	5.0	~~NaN~~ 3,25	1.0	4.0	3.0	− 3,25
B	2.0	3.0	5.0	~~NaN~~ 3,25	3.0	− 3,25
C	5.0	4.0	5.0	4.0	~~NaN~~ 4,5	− 4,5
D	3.0	2.0	~~NaN~~ 2,0	2.0	1.0	− 2,0

Normalize by 3			Normalize by user avg	
AB	-0.89		AB	-0.94
AC	0.105		AC	-0.21
AD	-0.13		AD	0.478
BC	0.282		BC	0.172
BD	0.0		BD	-0.32
CD	-0.25		CD	0.353

A	3.25
B	3.25
C	4.50
D	2.00

Examples- Jaccard

	M1	M2	M3	M4	M5	M6
A	5.0	NaN	1.0	4.0	NaN	NaN
B	2.0	3.0	5.0	NaN	1.0	NaN
C	4.0	4.0	NaN	4.0	NaN	3.0

	M1	M2	M3	M4	M5	M6
A	1	0	0	1	0	0
B	0	1	1	0	0	0
C	1	1	0	1	0	1

$$J(A,B) = \frac{S_A \cap S_B}{S_A \cup S_B}$$

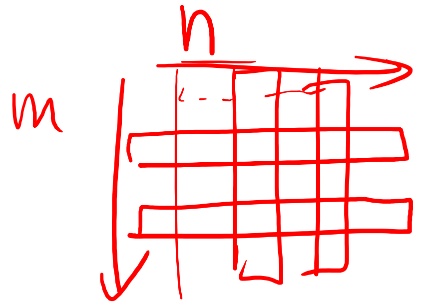
$$J(A,B) = 0$$

$$J(A,C) = \frac{2}{4} = 0.5$$

$$J(B,C) = \frac{1}{5} = 0.2$$

Things to consider in Recommender Systems

- Time complexity of operations



$m C_2 \times n$

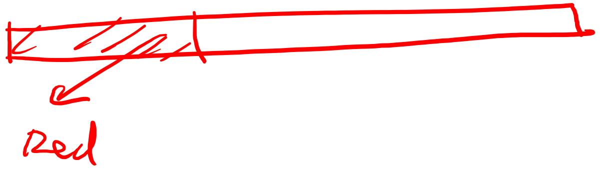
$\frac{m \times (m-1)}{2} \times n$

$\Rightarrow O(m^2 \cdot n)$

$O(n^2 \cdot m)$

- The data is often large scale

- The data is sparse



clustering

M_1	M_2	M_3
1	0	0
0	0	1
0	1	0

Recommender system in large scale

Where to place components?

- Example: Matrix Factorization
- Offline:
  - Collect sample of play data
  - Run batch learning algorithm to produce factorization
  - Publish item factors
- Nearline:
  - Solve user factors
  - Compute user-item products
  - Combine
- Online:
  - Presentation-context filtering
  - Serve recommendations

NETFLIX

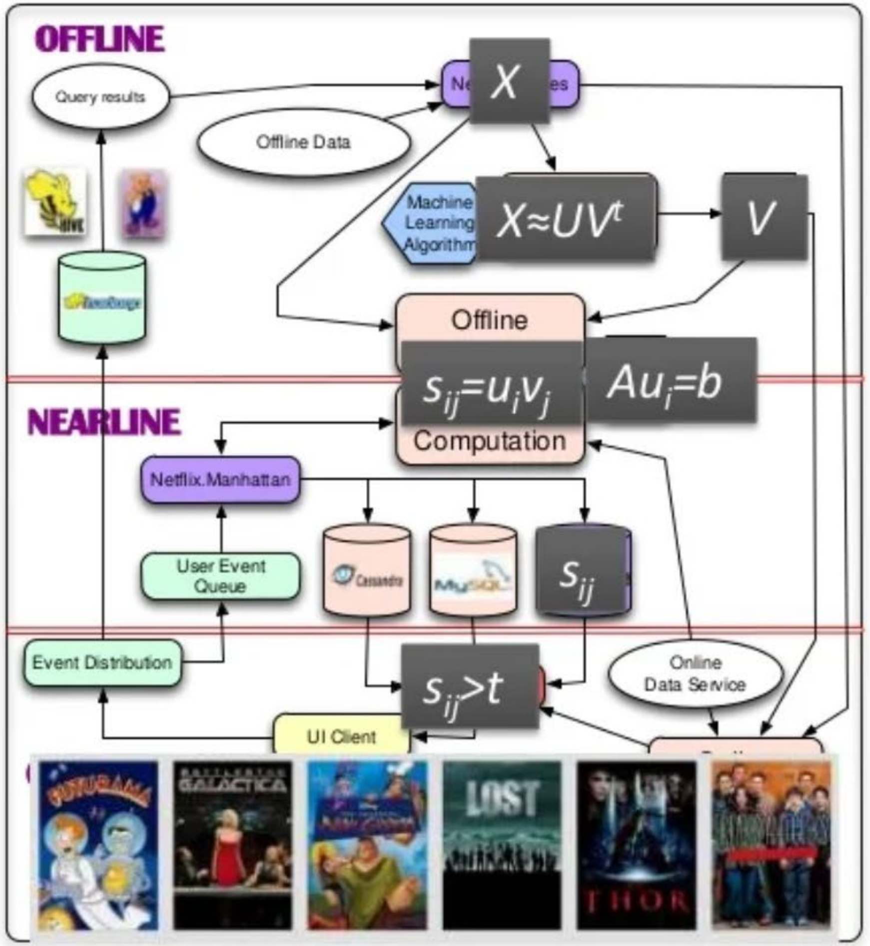


Image credit:
Justin Basilico
RecSys 2013