

Recommendation Systems

UE 141 Spring 2013

Jing Gao
SUNY Buffalo

Recommendation Systems

- **Data**

items

users

1	3	4			
	3	5			5
		4	5		5
		3			
		3			
2			2		2
				5	
	2	1			1
	3			3	
1					

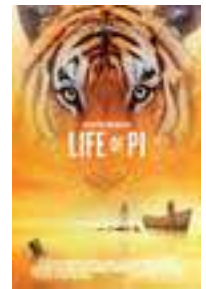
- **Goal**

- Learn what a user might be interested in and recommend other items he might like

Applications of Recommendation Systems

- **Amazon or other online shopping websites**
 - Huge amount of purchase data => Recommend other items to users
- **Movie/music/book/hotel/restaurant recommendation**
 - For example, Netflix, TripAdvisor, Yelp.....
 - Based on user ratings, recommend other items users might like
 - Attract and maintain users in the service
- **Recommend news, webpages, friends.....**

Movie Recommendation



User 1	1	1	1	0	0
User 2	0	0	0	1	1
User 3	0	1	0	0	1
User 4	1	1	1	0	?

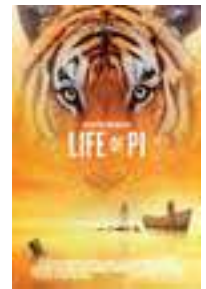
Interested: 1

Not interested: 0

What We Do in Real-life Recommendations?

- **Will I be interested in a new movie?**
 - I have a friend who share very similar taste with me—
Follow her recommendation of the new movie
 - I have a friend who has the “opposite” interest in movies—
Do not follow her recommendation of the new movie
 - I have a friend who sometime agrees with me but
sometime disagrees with me—I should not consider her
recommendation of the new movie seriously

Collaborative Filtering



User 1	1	1	1	0	0
User 2	0	0	0	1	1
User 3	0	1	0	0	1
User 4	1	1	1	0	?

User 1 & User 4: Positively correlated

User 2 & User 4: Negatively correlated

User 3 & User 4: Uncorrelated



$0 \Rightarrow 0$

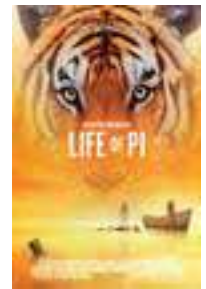
$1 \Rightarrow 0$

do not consider



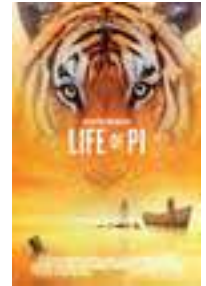
$? = 0$

Ratings from 1 to 5



User 1	5	4	3	1	3
User 2	1	1	2	5	4
User 3	2	5	1	2	5
User 4	5	4	4	1	?

Compute Average Rating



User 1	5	4	3	1	3
User 2	1	1	2	5	4
User 3	2	5	1	2	5
User 4	5	4	4	1	?

3.25

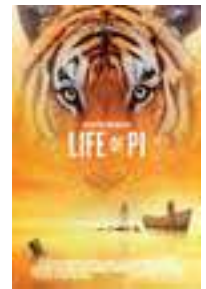
3.5

2.5

2.25

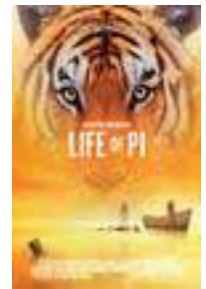
4

Subtract Average Rating



User 1	1.75	0.5	0.5	-1.25	-1
User 2	-2.25	-2.5	-0.5	2.75	0
User 3	-1.25	1.5	-1.5	-0.25	1
User 4	1.75	0.5	1.5	-1.25	?

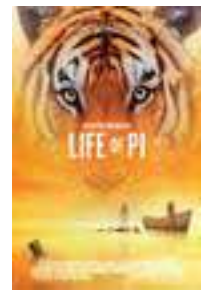
Computing Similarity



User 1	1.75	0.5	0.5	-1.25	-1
User 2	-2.25	-2.5	-0.5	2.75	0
User 3	-1.25	1.5	-1.5	-0.25	1
User 4	1.75	0.5	1.5	-1.25	?

Compute a similarity score between two users: The higher the score is, the more likely they enjoy the same movies

Pearson Correlation Measure

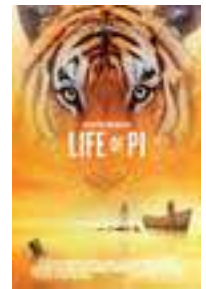


User 1	1.75	0.5	0.5	-1.25	-1
User 4	1.75	0.5	1.5	-1.25	?

$$\frac{1.75 \cdot 1.75 + 0.5 \cdot 0.5 + 0.5 \cdot 1.5 + (-1.25) \cdot (-1.25)}{\sqrt{1.75^2 + 0.5^2 + 0.5^2 + (-1.25)^2} \sqrt{1.75^2 + 0.5^2 + 1.5^2 + (-1.25)^2}}$$

$$= 0.9305$$

Computing New Rating



User 1	1.75	0.5	0.5	-1.25	-1
User 2	-2.25	-2.5	-0.5	2.75	0
User 3	-1.25	1.5	-1.5	-0.25	1
User 4	1.75	0.5	1.5	-1.25	?

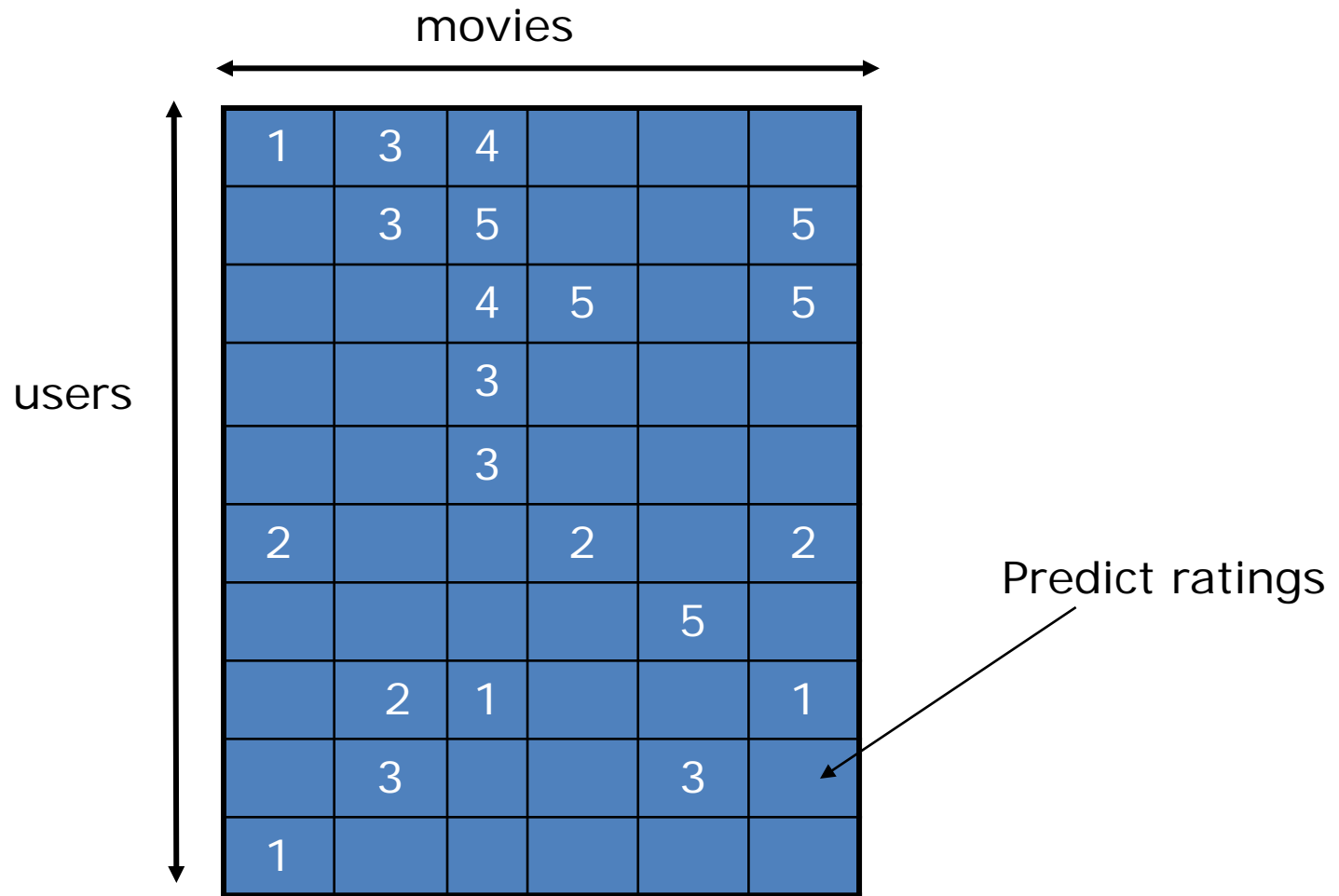
User 1 & User 4: 0.9305 User 2 & User 4: -0.7904

User 3 & User 4: -0.4382



$$? = 4 + 0.9305 * (-1) + (-0.7904) * 0 + (-0.4382) * 1 = 2.6313$$

Ratings Data



Modification

- **Huge sparse data**
 - Compute average ratings based on the available ratings
 - Compute similarity between two users based on the ratings on the items that both users give ratings to
 - Compute the weighted average only upon top K users who share the most similar ratings with the target user

Mathematical Formulation

$r_{i,k}$ = rating of user i on item k

I_k = the set of users who give ratings to item k

Average rating for item k is

$$\bar{m}_k = \frac{1}{|I_k|} \sum_{m \in I_k} r_{m,k}$$

Predicted vote for user i on item k is a weighted sum

$$r_{i,k} = \bar{m}_k + \sum_{m=1}^K \hat{a}_{i,m} (r_{m,k} - \bar{m}_k)$$

User-centered Recommendation

$r_{i,k}$ = rating of user i on item k

I_i = items for which user i has generated a rating

Average rating for user i is

$$m_i = \frac{1}{|I_i|} \sum_{j \in I_i} r_{i,j}$$

Predicted vote for user i on item j is a weighted sum

$$r_{i,k} = m_i + \sum_{j=1}^K w_{k,j} (r_{i,j} - m_j)$$

Question

- How does Facebook recommend possible friends for you?