



Carnegie Mellon University
Tepper School of Business

46-886: Machine Learning Fundamentals

Amr Farahat

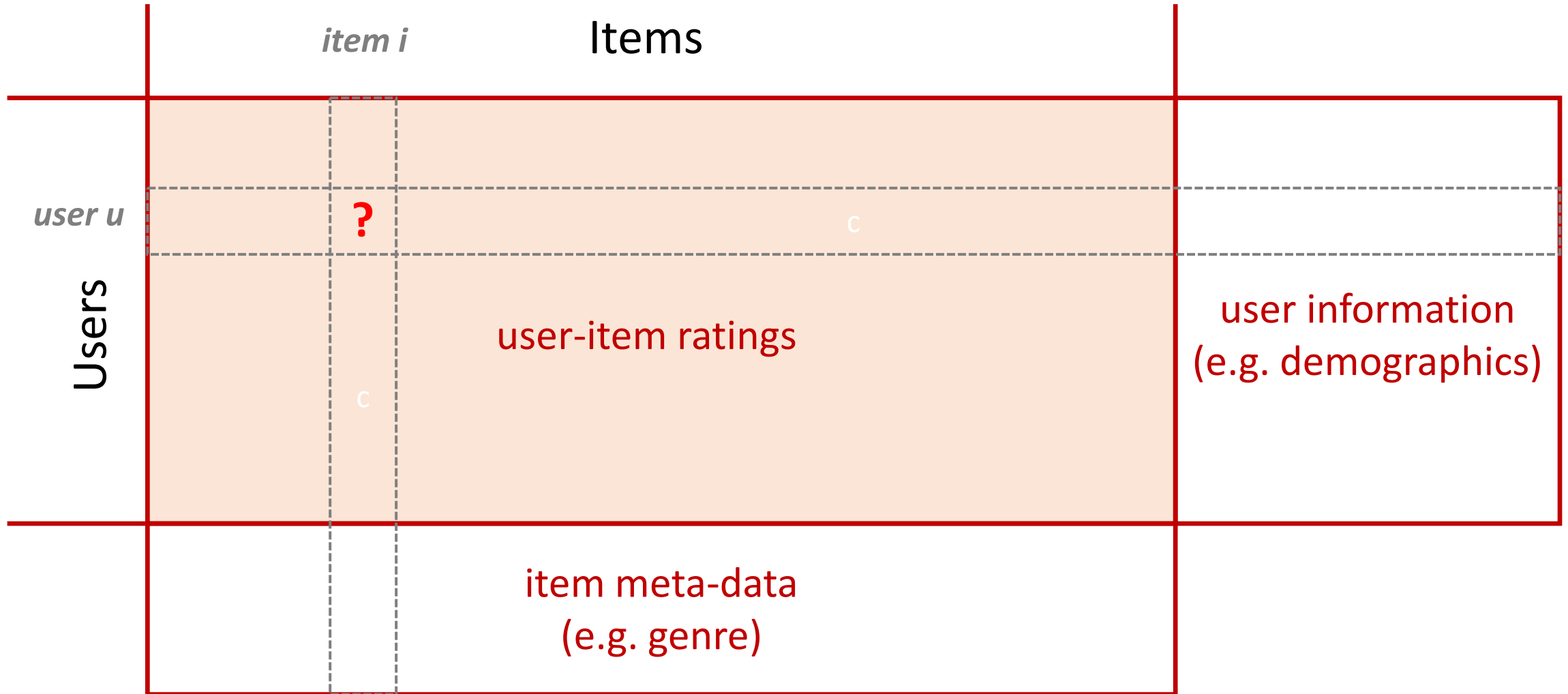
Recommender Systems: Neighborhood-Based Collaborative Filtering

Much of this slide deck is derived/borrowed from course material
I've co-taught at MIT



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Collaborative Filtering



- For user u and item i , we want to predict how user u will rate item i

Collaborative Filtering Systems

- Two broad types of CF methods:
 - Neighborhood CF
 - User-User
 - Item-Item
 - Model-based CF

Neighborhood CF: User-User

- User-User CF:
 - Identify users who are ‘very similar’ to user u (based on ratings) and who have actually rated i
 - Prediction: Average those users’ ratings, using user-user similarity as the weight
- Examples of “similarity” metrics
 - Correlation between rating vectors
 - “Distance” between rating vectors

Neighborhood CF: User-User

- Example: We want to predict Al's rating for M2

	M1	M2	M3	M4	M5
<i>Al</i>	5	?	4	5	?
<i>Ben</i>	4	1	?	4	2
<i>Carrie</i>	1	4	5	2	4
<i>Debra</i>	?	?	?	4	3

- Basic idea: Average M2's ratings from the other users weighted by their similarity to Al.

Neighborhood CF: Item-Item

- Item-based CF = user-based CF “rotated by 90 degrees”

	M1	M2	M3	M4	M5
<i>Al</i>	5	?	4	5	?
<i>Ben</i>	4	1	?	4	2
<i>Carrie</i>	1	4	5	2	4
<i>Debra</i>	?	?	?	4	3

- Basic idea: Average Al's ratings for other movies weighted by their similarity to M2.
- Item-Item CF:
 - Identify items that are ‘very similar’ to item i and which have been rated by user u
 - Prediction: Average those items' ratings, using item-item similarity as the weight