Carnegie Mellon University
Tepper School of Business

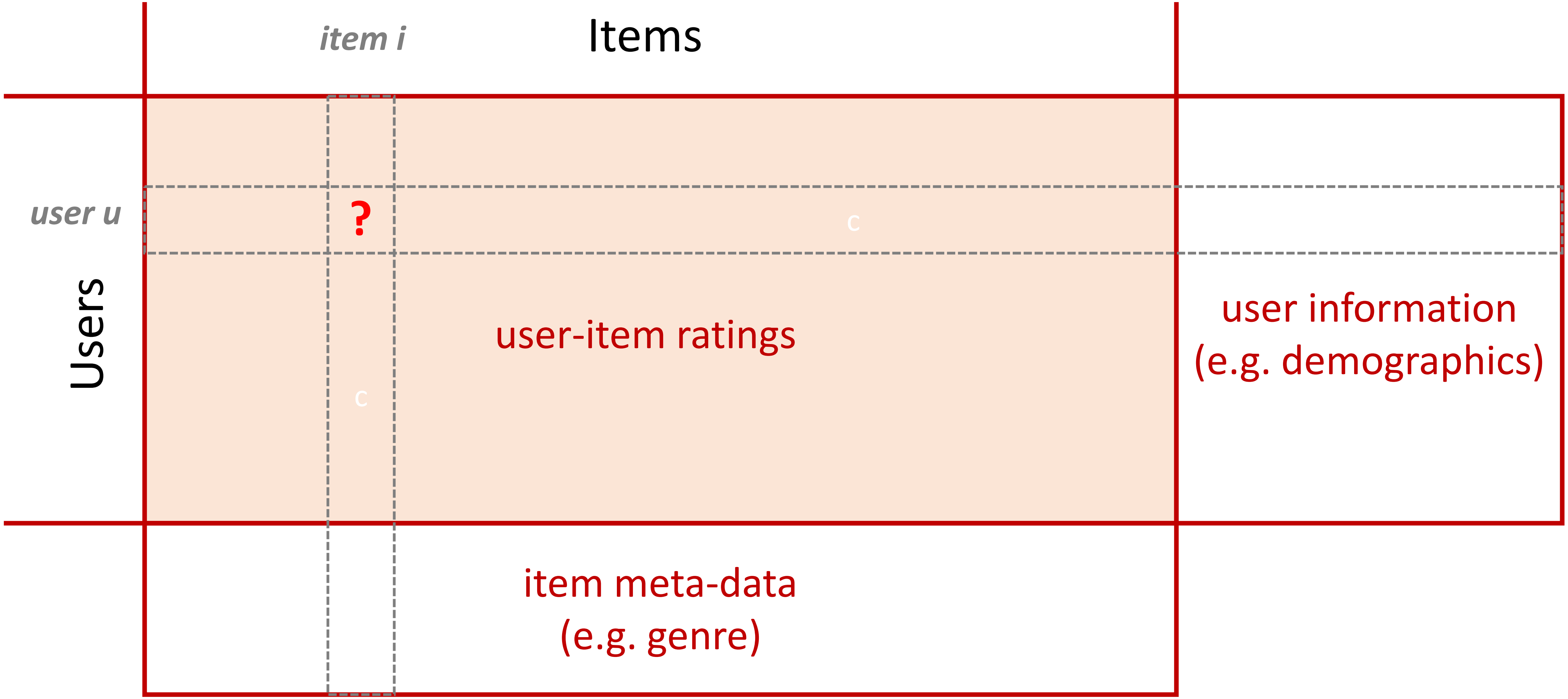
46-886: Machine Learning Fundamentals

Amr Farahat

Recommender Systems: Model-Based Collaborative Filtering

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

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

Model-Based Collaborative Filtering

Notation

User	Movie																										
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	…
John				5					1				3			4											…
Isabel														4				4					1		2		…
Charles			3				2				2				5												…
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…

- Index users by $u = 1, \dots, 6040$ and movies by $m = 1, \dots, 3900$
- OBS: collection of the **observed** user-movie pairs (u, m)
  - (John , D) , (John , I) , (John , M) , (John , P) , (Isabel , N) , …
- $OBSR_{u,m}$: observed rating of the user-movie pair (u, m)
  - For example $OBSR_{John,D} = 5$

Some Intuition

- More generally, how can we estimate missing user ratings?

User	Movie																									
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	…
John	5		3		1		2		4																		…
Isabel	5		3		1		2		?																		…
Charles											2		1		2		1		1								…
Paul											3		3		2		3		?								…
Rahul											4		5		2		5		3								…
Jose																					1	5	2	2	4		…
Caroline																					5	1	4	4	?		…
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…

Where's the structure?

User	Movie																									
	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	…
John				5					1				3			4											…
Isabel														4				4					1		2		…
Charles			3				2				2				5												…
Paul					1				1										4		5						…
Rahul												2												4			…
Jose		5				3											2			2						1	…
Caroline								4							1							4					…
Stephanie	3				2					2	2							3						5			…
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…

The previous slides were artificial of course: there were visibly obvious ways to impute the missing ratings using particular users or movies. But even so:

- How could we extrapolate to other movies not viewed by any users in a cluster?
- How do we pool information across users and movies in a more systematic way?

New Concept:
Archetype Users

Concept of Archetypal Users

- "Archetypal user": idealized stereotype of movie-watchers:
  - Archetype 1: ("Culture Gourmand") Enjoys movies on topics that allow for introspection and adventure or nostalgic trips of childlike awe and wonderment.
  - Archetype 2: ("Artistic Activist") Sees movies as both reflective of society and implements for social change. Also is a visual aesthete.
  - …
  - Archetype 9: ("Emotional Escapist") Prefers movies that engage in the emotional side of life, via tears, laughter, etc., giving the STEM side of the brain a few hours of rest.
- We will create a set of k archetypal users (we will use $k = 9$)

Archetype Users are characterized by their ratings of all of the movies

Archetype	Hillbilly Elegy	The Dig	Lost Girls	Mank	The Midnight Sky	I Care a Lot	Rebecca	Moxie	The King	…
1	3	5	2	5	5	3	3	4	1	…
2	5	4	5	2	3	1	5	4	5	…
…	…	…	…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…	…	…	…
9	2	5	3	5	4	5	1	4	3	…

- Index archetypes by $a = 1, \ldots, 9$ and movies by $m = 1, \ldots, 3900$
- $S_{a,m}$: supposed rating of movie m by archetype a
  - For example, in the above table: $S_{2,Mank} = 2$ and $S_{9,Mank} = 5$

Concept of Archetypal Users, cont.

- We could specify the archetypes, in which case they might be described as in the previous slide, and we could then model each user as a weighted linear combination of the archetypes

- Or we could leave it up to the model to determine the archetypes. This will yield a more accurate model, but the model might not necessarily be so interpretable
  - Kind of like in clustering where the model determines the clusters and we then have to try to interpret the cluster means afterwards

- We will presume that there are k archetypal users, where k is a small number (we will use $k = 9$)

Archetypes Ratings of Movies

Suppose the archetype ratings of movies are:

Archetype	Movie																										
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	…
1	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	3	…	…	…	…	…	…	…	…	…	…
2	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	2	…	…	…	…	…	…	…	…	…	…
3	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	4	…	…	…	…	…	…	…	…	…	…
4	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	5	…	…	…	…	…	…	…	…	…	…
5	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	3	…	…	…	…	…	…	…	…	…	…
6	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	2	…	…	…	…	…	…	…	…	…	…
7	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	1	…	…	…	…	…	…	…	…	…	…
8	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	4	…	…	…	…	…	…	…	…	…	…
9	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	3	…	…	…	…	…	…	…	…	…	…

Concept of each User as a weighted combination of archetypes

- **Idea**: each user can be approximated as some linear weighted combination of archetypes, perhaps:
  - John ~ **0.7** x (Arch.1) + **0.1** x (Arch.2) + … + **0.2** x (Arch.9)
  - Isabel ~ **0.8** x (Arch.1) + **1.6** x (Arch.2) + … **– 0.3** x (Arch.9)
  - Charles ~ **0.2** x (Arch.1) **– 0.6** x (Arch.2) + … + **1.1** x (Arch.9)
- Each user is characterized by $k = 9$ weights, one for each archetype
  - User u ~ $\boldsymbol{w_{u,1}}$ x (Arch.1) + $\boldsymbol{w_{u,2}}$ x (Arch.2) + … + $\boldsymbol{w_{u,9}}$ x (Arch.9)
  - e.g., $w_{Charles,1} = \mathbf{0.2}$, $w_{Charles,2} = \mathbf{-0.6}$, …, $w_{Charles,9} = \mathbf{1.1}$

→ There are 9 x 6,040 = 54,360 weights – one for each user-archetype pair. We will determine/compute the weights that best fit the training data (just like in our other prediction models).

Users' composition weights of Archetypes are to be determined by the model fitting procedure

Let us now suppose that the user weights of archetypes are to be determined by the model:

User	Archetype								
	1	2	3	4	5	6	7	8	9
John	…	…	…	…	…	…	…	…	…
Isabel									
Charles									
Paul									
Rahul	…	…	…	…	…	…	…	…	…
Jose	$w_{Jose,1}$	$w_{Jose,2}$	$w_{Jose,3}$	$w_{Jose,4}$	$w_{Jose,5}$	$w_{Jose,6}$	$w_{Jose,7}$	$w_{Jose,8}$	$w_{Jose,9}$
Caroline	…	…	…	…	…	…	…	…	…
Stephanie									
…	…	…	…	…	…	…	…	…	…

Archetype Ratings of Movies are to be determined by the model fitting procedure

Let us also suppose the archetype ratings of movies are to be solved by the model:

Archetype	Movie																									
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	…
1																…	$s_{1,Q}$	…									…
2																…	$s_{2,Q}$	…									…
3																…	$s_{3,Q}$	…									…
4																…	$s_{4,Q}$	…									…
5																…	$s_{5,Q}$	…									…
6																…	$s_{6,Q}$	…									…
7																…	$s_{7,Q}$	…									…
8																…	$s_{8,Q}$	…									…
9	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	$s_{9,Q}$	…	…	…	…	…	…	…	…	…	…

Formula for Jose's predicted rating of movie Q

User	Archetype								
	1	2	3	4	5	6	7	8	9
John	…	…	…	…	…	…	…	…	…
Isabel									
Charles									
Paul									
Rahul	…	…	…	…	…	…	…	…	…
Jose	$w_{Jose,1}$	$w_{Jose,2}$	$w_{Jose,3}$	$w_{Jose,4}$	$w_{Jose,5}$	$w_{Jose,6}$	$w_{Jose,7}$	$w_{Jose,8}$	$w_{Jose,9}$
Caroline	…	…	…	…	…	…	…	…	…
Stephanie									
…	…	…	…	…	…	…	…	…	…

Archetype	Movie																										
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	…
1																…	$s_{1,Q}$	…									…
2																…	$s_{2,Q}$	…									…
3																…	$s_{3,Q}$	…									…
4																…	$s_{4,Q}$	…									…
5																…	$s_{5,Q}$	…									…
6																…	$s_{6,Q}$	…									…
7																…	$s_{7,Q}$	…									…
8																…	$s_{8,Q}$	…									…
9	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	$s_{9,Q}$	…	…	…	…	…	…	…	…	…	…

$$R_{\text{Jose,Q}} = w_{\text{Jose,1}} \cdot S_{1,Q} + w_{\text{Jose,2}} \cdot S_{2,Q} + \cdots + w_{\text{Jose,9}} \cdot S_{9,Q}$$

Illustration of model computation: prediction of Jose's rating of movie Q

- Let us illustrate how we compute Jose's rating of movie Q using the illustrative data numbers shown…

Users’ composition weights of Archetypes

Suppose the user composition weights of archetypes are:

User	Archetype								
	1	2	3	4	5	6	7	8	9
John	…	…	…	…	…	…	…	…	…
Isabel									
Charles									
Paul									
Rahul	…	…	…	…	…	…	…	…	…
Jose	.2	.3	.1	.9	-.3	-.4	.8	-.8	.7
Caroline	…	…	…	…	…	…	…	…	…
Stephanie									
…	…	…	…	…	…	…	…	…	…

Archetype Ratings of Movies

Suppose the archetype ratings of movies are:

Archetype	Movie																										
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	…
1																…	3	…									…
2																…	2	…									…
3																…	4	…									…
4																…	5	…									…
5																…	3	…									…
6																…	2	…									…
7																…	1	…									…
8																…	4	…									…
9	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	3	…	…	…	…	…	…	…	…	…	…

Jose's predicted rating of movie Q

User	Archetype								
	1	2	3	4	5	6	7	8	9
John	…	…	…	…	…	…	…	…	…
Isabel									
Charles									
Paul									
Rahul	…	…	…	…	…	…	…	…	…
Jose	.2	.3	.1	.9	-.3	-.4	.8	-.8	.7
Caroline	…	…	…	…	…	…	…	…	…
Stephanie									
…	…	…	…	…	…	…	…	…	…

Archetype	Movie																										
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	…
1																…	3	…									…
2																…	2	…									…
3																…	4	…									…
4																…	5	…									…
5																…	3	…									…
6																…	2	…									…
7																…	1	…									…
8																…	4	…									…
9	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	3	…	…	…	…	…	…	…	…	…	…

$$R_{\text{Jose, Q}} = .2\times 3 + .3\times 2 + .1\times 4 + .9\times 5 - .3\times 3 - .4\times 2 + .8\times 1 - .8\times 4 + .7\times 3 = 4.1$$

Jose’s observed rating of movie Q

User	Movie																									
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	…
John				5					1				3			4											…
Isabel														4				4					1		2		…
Charles			3				2				2				5												…
Paul					1				1										4		5						…
Rahul												2												4			…
Jose		5				3											2			2						1	…
Caroline								4							1							4					…
Stephanie	3				2					2	2							3						5			…
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…