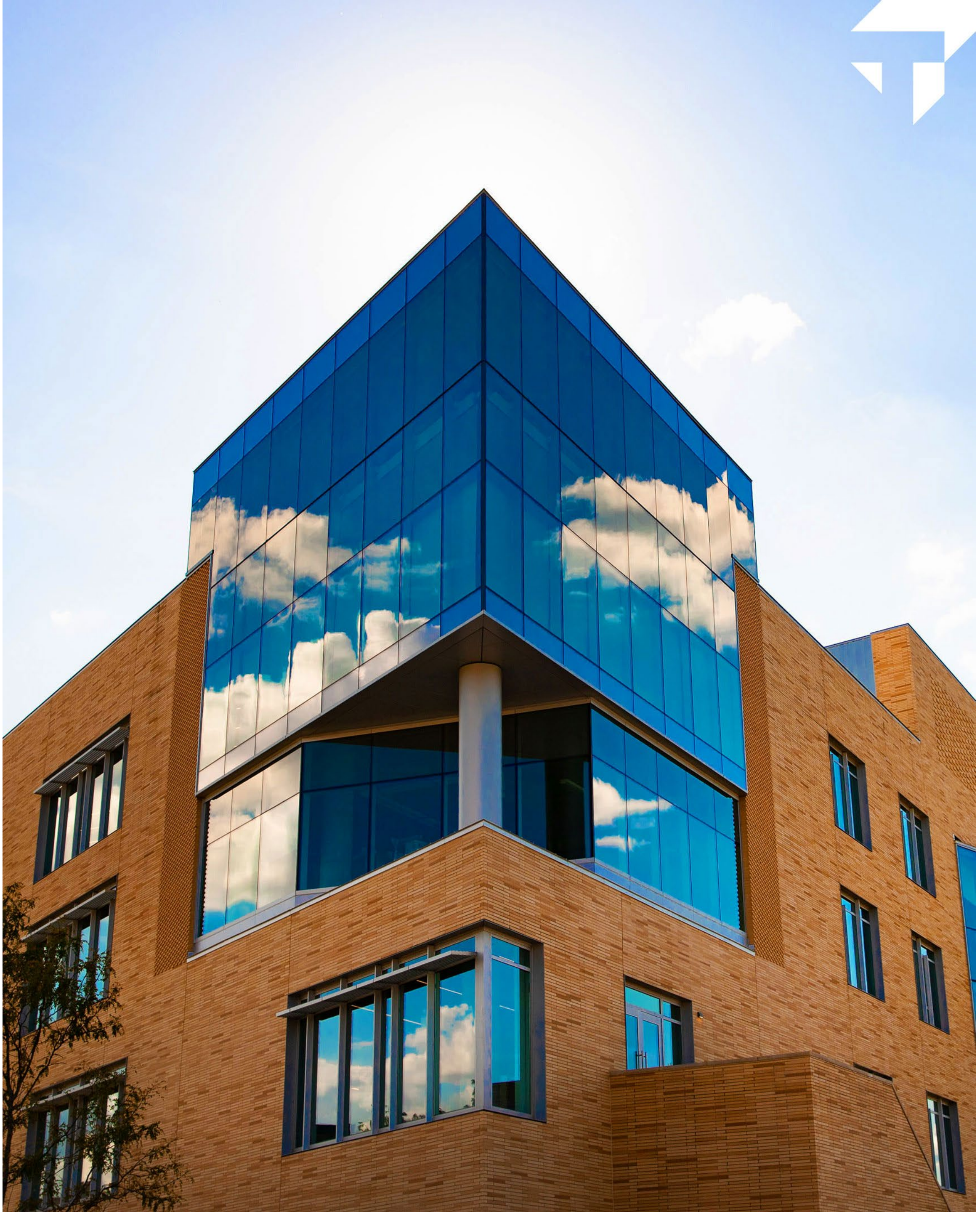

Carnegie Mellon University
Tepper School of Business

46-886: Machine Learning Fundamentals

Amr Farahat

Recommender Systems:
Model-Based Collaborative Filtering – Part 2

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

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	…	…	…	…	…	…	…	…	…	…

$$R_{\text{Jose, Q}} = .2\times3+.3\times2+.1\times4+.9\times5-.3\times3-.4\times2+.8\times1-.8\times4+.7\times3 = 4.1$$

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			…
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…

Residual for model’s predicted value for Jose of movie Q is

$$OBSR_{Jose,Q} - 4.1 = 2 - 4.1 = \mathbf{-2.1}$$

Least Squares Fitting

- Question: Which objects do we need to determine?
  - $S_{a,m}$: supposed rating of movie m by archetype a
  - $w_{u,a}$: weight of archetype a in the linear combination of user u
- Let us compute these values to minimize the sum of squared residuals, namely the difference between observed and predicted user-movie ratings for all of the user-movie pairs in the observed ratings

$$\text{minimize} \sum_{\substack{\text{all observed} \\ \text{user-movie} \\ \text{pairs } (u,m)}} \left(\text{OBSR}_{u,m} - \left(w_{u,1}S_{1,m} + w_{u,2}S_{2,m} + \cdots + w_{u,9}S_{9,m}\right)\right)^2$$

Summary of Model-Based CF

1. Observed user-movie 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					1				3			4											…
Isabel														4				4					1		2		…
Charles			3				2				2				5												…
Paul					1														4		5						…
Rahul												2												4			…
Jose		5				3											2			2						1	…
Caroline								4							1							4					…
Stephanie	3				2					2	2							3						5			…
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…

$OBSR_{u,m}$

2. Archetype-movie ratings

Archetype	Hillbilly Elegy	The Dig	irls	k	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	…

$S_{a,m}$

3. User-archetype weights

User	Archetype								
	1	2	3	4	5	6	7	8	9
John	.7	.1	…	…	…	…	…	…	.2
Isabel			…	…	…	…	…	…	-.3
Charles			…	…	…	…	…	…	1.1
Paul	…	…	…	…	…	…	…	…	…
Rahul	…	…	…	…	…	…	…	…	…
Jose	**.2**	**.3**	**.1**	**.9**	**-.3**	**-.4**	**.8**	**-.8**	**.7**
Caroline	…	…	…	…	…	…	…	…	…
Stephanie	…	…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…	…	…

$w_{u,a}$

4. Least Squares Fitting

$$\text{minimize} \sum_{\substack{\text{all observed} \\ \text{user-movie} \\ \text{pairs } (u,m)}} \left(\mathrm{OBSR}_{u,m} - (w_{u,1}S_{1,m} + w_{u,2}S_{2,m} + \cdots + w_{u,9}S_{9,m})\right)^2$$

Model-Based Collaborative Filtering

- This modeling approach is an example of “model-based collaborative filtering”
  - collaborative: leverages other users’ ratings to predict unobserved ones
  - filtering: of all the choices, filter the ones that’s best for the individual
- We specify the number of archetypes k, and the model determines the archetypes and the user-archetype weights
  - This approach shares similarities with clustering in the sense that it produces customer archetypes that can be easy to “interpret”
- Active area of research and development