Slide10

Sunday, July 18, 2021 5:35 PM

TutorialSlide10

Xiaochen Yang
xiaochen.yang@glasgow.ac.uk

This week's content

- Content-based filtering systems
- Collaborative filtering systems
  - Neighbourhood-based collaborative filtering
  - Model-based collaborative filtering (UV-decomposition)
- Evaluation measures

Recommender systems

- Aim: predict the rating or preference a user would give to an item.
- Data: **utility matrix**

		Items					
		1	2	...	i	...	R
	1	5			3		4
	2		2			3	4
Users	⋮						
	u		2			5	
	⋮						
	n	2		3	4	1	
(to predict) active user	a		1		?	3	

- numerical ratings / unary ratings
  1–10 purchased/not purchased 1 / 0 inferred from user

Content-based filtering

Idea: compare representations of content describing an item to representations of content that interests the user

1. construct item profile
2. construct user profile
3. compute the similarity between item and user profiles

Content-based filtering (example)

	Harry Potter	Star Wars	Titanic	Love Actually
Alice	3	2	4	
Bob		5		
Carol			4	3
David	3		1	

1. construct item profile

feature to describe movie

	Fantasy	Romantic
Harry Potter	1	0
Star Wars	1	0
Titanic	0	1
Love Actually	0	1

Content-based filtering (example)

item \ feature	Fantasy	Romantic
Harry Potter	1	0
Star Wars	1	0
Titanic	0	1
Love Actually	0	1

2. construct item profile: for numerical ratings, normalise the utilities by subtracting the average rating for a user need to match w/ item profile

	HP	SW	Titanic	LA	Fantasy	Romantic
Alice	3	2	4		−0.5	1

Alice's average rating: $\frac{3+2+4}{3} = 3$ normalization

Alice's profile on Fantasy: $\frac{(3-3)+(2-3)}{2} = -0.5$ ⇒ compute cos similarity

Alice's profile on Romantic: $\frac{4-3}{1} = 1$

Neighbourhood-based collaborative filtering

(2)

Idea: recommendation for an active user a is made by looking at the users that are most similar to a and recommending items that these users like.

1. calculate the similarity from a to all other users u
2. select k users most similar to a who have rated item i
3. make predictions for user a on item i

Neighbourhood-based collaborative filtering

1. calculate the similarity from a to all other users u

- Pearson correlation coefficient

$$\text{sim}(a,u) = \frac{\sum_{i\in I}(r_{a,i}-\bar{r}_a)(r_{u,i}-\bar{r}_u)}{\sqrt{\sum_{i\in I}(r_{a,i}-\bar{r}_a)^2}\sqrt{\sum_{i\in I}(r_{u,i}-\bar{r}_u)^2}}$$

- Cosine similarity

$$\text{sim}(a,u) = \frac{\boldsymbol{r}_a \cdot \boldsymbol{r}_u}{\|\boldsymbol{r}_a\| \cdot \|\boldsymbol{r}_u\|}$$

- Jaccard similarity measure (only 0 or 1)

$$\text{sim}(a,u) = \frac{|I_a \cap I_u|}{|I_a \cup I_u|}$$

Neighbourhood-based collaborative filtering

2. select k users most similar to a who have rated item i
3. make predictions for user a on item i

- Let N be the set of k users most similar to a who have rated item i
- Prediction for item i of user a:
  - $p_{a,i} = \frac{1}{|N|}\sum_{u\in N} r_{u,i}$ simple average
  - $p_{a,i} = \frac{\sum_{u\in N} r_{u,i}\cdot \text{sim}(a,u)}{\sum_{u\in N}|\text{sim}(a,u)|}$ weighted avege, if a, u similar ⇒ higher influence
  - $p_{a,i} = \bar{r}_a + \frac{\sum_{u\in N}(r_{u,i}-\bar{r}_u)\times \text{sim}(a,u)}{\sum_{u\in N}|\text{sim}(a,u)|}$ delete bias

Collaborative filtering (example)

user \ item	i1	i2	i3	i4	i5	i6	avg for user i
u1	.	4	3	4	3	3	3.40
u2	2	4	.	4	2	3	3.00
u3	4	5	.	.	.	1	3.33
u4	3	5	.	4	5	1	3.60
u5	5	1	3	3	4	1	2.83
u6	5	2	2	3	3	1	2.67
u7	.	.	5	1	.	4	3.33
u8	3	2	1	3	.	3	2.40
u9	5	1	5	5	3	.	3.80
u10	5	3	2	3	2	3	3.00

Collaborative filtering (example)

	i1	i2	i3	i4	i5	i6	avg
u1	.	4	3	4	3	3	3.40
u2	2	4	.	4	2	3	3.00
u3	4	5	.	.	.	1	3.33
u4	3	5	.	4	5	1	3.60
u5	5	1	3	3	4	1	2.83
u6	5	2	2	3	3	1	2.67
u7	.	.	5	1	.	4	3.33
u8	3	2	1	3	.	3	2.40
u9	5	1	5	5	3	.	3.80
u10	5	3	2	3	2	3	3.00

2,4,5,6 rated for both 1, 2

Pearson correlation between users 1 and 2

$$\frac{(4-3.4)\cdot(4-3)+(4-3.4)\cdot(4-3)+(3-3.4)\cdot(2-3)+(3-3.4)\cdot(3-3)}{\sqrt{2\times(4-3.4)^2+2\times(3-3.4)^2}\sqrt{2\times(4-3)^2+(2-3)^2+(3-3)^2}}$$

$= 0.906$

Collaborative filtering (example)

	i1	i2	i3	i4	i5	i6	avg
u1	.	4	3	4	3	3	3.40
u2	2	4	.	4	2	3	3.00
u3	4	5	.	.	.	1	3.33
u4	3	5	.	4	5	1	3.60
u5	5	1	3	3	4	1	2.83
u6	5	2	2	3	3	1	2.67
u7	.	.	5	1	.	4	3.33
u8	3	2	1	3	.	3	2.40
u9	5	1	5	5	3	.	3.80
u10	5	3	2	3	2	3	3.00

Cosine similarity between users 1 and 2 (missing is 0, for all values)

$$\frac{0\cdot2+4\cdot4+3\cdot0+4\cdot4+3\cdot2+3\cdot3}{\sqrt{0^2+4^2+3^2+4^2+3^2+3^2}\sqrt{2^2+4^2+0^2+4^2+2^2+3^2}}$$

$= 0.874$

Collaborative filtering (example)

	i1	i2	i3	i4	i5	i6	avg
u1	.	4	3	4	3	3	3.40
u2	2	4	.	4	2	3	3.00
u3	4	5	.	.	.	1	3.33
u4	3	5	.	4	5	1	3.60
u5	5	1	3	3	4	1	2.83
u6	5	2	2	3	3	1	2.67
u7	.	.	5	1	.	4	3.33
u8	3	2	1	3	.	3	2.40
u9	5	1	5	5	3	.	3.80
u10	5	3	2	3	2	3	3.00

Jaccard similarity between users 1 and 2

$\frac{\text{rated for both}}{\text{rated for at least one user}} = \frac{4}{6} = 0.667$

all have been rated by one of the users

Collaborative filtering (example)

prediction: for user 1, item 1:

	i1	i2	i3	i4	i5	i6	Pearson
u1	.	4	3	4	3	3	1
u2	2	4	.	4	2	3	0.905 ✓
u3	4	5	.	.	.	1	1 ✓
u4	3	5	.	4	5	1	0.457
u5	5	1	3	3	4	1	-0.272
u6	5	2	2	3	3	1	0.327
u7	.	.	5	1	.	4	-0.971
u8	3	2	1	3	.	3	0.302
u9	5	1	5	5	3	.	-0.302
u10	5	3	2	3	2	3	0.667 ✓

Using $k = 3$, i.e. users 3, 2 and 10

simple average: $r_{11} = \frac{2+4+5}{3} = 3.67$

Collaborative filtering (example)

	i1	i2	i3	i4	i5	i6	Pearson
u1	.	4	3	4	3	3	1
u2	2	4	.	4	2	3	0.905
u3	4	5	.	.	.	1	1
u4	3	5	.	4	5	1	0.457
u5	5	1	3	3	4	1	-0.272
u6	5	2	2	3	3	1	0.327
u7	.	.	5	1	.	4	-0.971
u8	3	2	1	3	.	3	0.302
u9	5	1	5	5	3	.	-0.302
u10	5	3	2	3	2	3	0.667

Using $k = 3$, i.e. users 3, 2 and 10

weighted average: $r_{11} = \frac{2\cdot0.905+4\cdot1+5\cdot0.667}{1+0.905+0.667} = 3.56$

Collaborative filtering (example)

	i1	i2	i3	i4	i5	i6	avg	Pearson
u1	.	4	3	4	3	3	3.40	1
u2	2	4	.	4	2	3	3.00	0.905 ✓
u3	4	5	.	.	.	1	3.33	1 ✓
u4	3	5	.	4	5	1	3.60	0.457
u5	5	1	3	3	4	1	2.83	-0.272
u6	5	2	2	3	3	1	2.67	0.327
u7	.	.	5	1	.	4	3.33	-0.971
u8	3	2	1	3	.	3	2.40	0.302
u9	5	1	5	5	3	.	3.80	-0.302
u10	5	3	2	3	2	3	3.00	0.667 ✓

Using $k = 3$, i.e. users 3, 2 and 10

weighted average (normalised): $r_{11} =$

$$3.4 + \frac{(2-3)\cdot0.905+(4-3.33)\cdot1+(5-3)\cdot0.667}{1+0.905+0.667} = 3.83$$

avg. of user 1

UV-decomposition

(3) Idea: decompose the utility matrix M as UV, optimise U and V such that UV closely approximates M in the nonblank entries, and use the entry in UV to estimate the blank entries

$$\begin{bmatrix}5&2&4&4&3\\3&1&2&4&1\\2&&3&1&4\\2&5&4&3&5\\4&4&5&4&\end{bmatrix} = \begin{bmatrix}u_{11}&u_{12}\\u_{21}&u_{22}\\u_{31}&u_{32}\\u_{41}&u_{42}\\u_{51}&u_{52}\end{bmatrix} \times \begin{bmatrix}v_{11}&v_{12}&v_{13}&v_{14}&v_{15}\\v_{21}&v_{22}&v_{23}&v_{24}&v_{25}\end{bmatrix}$$

UV-decomposition

Idea: decompose the utility matrix M as UV, optimise U and V such that UV closely approximates M in the nonblank entries, and use the entry in UV to estimate the blank entries

$$\begin{bmatrix}5&2&4&4&3\\3&1&2&4&1\\2&&3&1&4\\2&5&4&3&5\\4&4&5&4&\end{bmatrix} = \underbrace{\begin{bmatrix}u_{11}&u_{12}\\u_{21}&u_{22}\\u_{31}&u_{32}\\u_{41}&u_{42}\\u_{51}&u_{52}\end{bmatrix}}_{5\times2} \times \underbrace{\begin{bmatrix}v_{11}&v_{12}&v_{13}&v_{14}&v_{15}\\v_{21}&v_{22}&v_{23}&v_{24}&v_{25}\end{bmatrix}}_{2\times5}$$

1. determine the number of latent dimensions size of matrix U, V
2. incrementally optimise the values in U and V by minimising the sum of squared error (one by one) $u_{11}, v_{11}, u_{12}, v_{12}, \ldots$

$$\text{SSE} = \sum_{(u,i)\in R}(r_{u,i} - \boldsymbol{u}_u \cdot \boldsymbol{v}_i)^2$$

(given rating − predicted predicted)²

3. predict the blank entries

Content-based filtering *vs* collaborative filtering

Content-based filtering

Pros

- Able to recommend new & unpopular items (feature profile)
- Able to provide explanations

Cons

- Finding the appropriate features is hard
- Cannot make recommendations for new users (no user profile)
- Overspecialisation: never recommends items outside user's content profile

Content-based filtering *vs* collaborative filtering

Collaborative filtering

Pros

- Works for any kind of item

Cons

- Hard to find a set of similar users when the utility matrix is very sparse (many 0's)
- Cannot make recommendations for new users
- Cannot recommend an item that has not been previously rated
- Popularity bias: tends to recommend popular items

Evaluation

- Evaluate predictions with known ratings (regression setting)
  - root mean squared error: $\text{RMSE} = \sqrt{\frac{\sum_{(u,i)\in T}(p_{u,i}-r_{u,i})^2}{|T|}}$
  - mean absolute error: $\text{MAE} = \frac{\sum_{(u,i)\in T}|p_{u,i}-r_{u,i}|}{|T|}$
- Evaluate Top-N recommendations: 0/1 model
  - Accuracy, precision (positive prediction rate), ROC (for classification setting)
- Evaluate the ranking of Top-N recommendation
  - Spearman's rank correlation between system's and user's (complete) rankings

Errata

- Example 2 (page 5)

$$p_{1,2} = 3.33 + \frac{(5-4.67)\times1+(2-3)\times0}{1+0} = 3.66.$$

- Page 7: definition of sum of squared error

$$\text{SSE} = \sum_{(u,i)\in R}(r_{u,i} - \boldsymbol{u}_u \cdot \boldsymbol{v}_i)^2,$$

- Example 4 (page 7)

$$(5-(x+1))^2+(2-(x+1))^2+(4-(x+1))^2+(4-(x+1))^2+(3-(x+1))^2.$$

We want the value of x that minimises the above sum, so we take the derivative and set it to 0, as:

$$-2\times((4-x)+(1-x)+(3-x)+(3-x)+(2-x)) = 0,$$

Tutorial sheet Q5

Three computers, A, B and C, have the numerical features listed below:

Feature \ item	A	B	C
Processor speed	3.06	2.68	2.92
Disk size	500	320	640
Main-memory size	6	4	6

A certain user has rated the three computers as follows: A: 4 stars, B: 2 stars, C: 5 stars.

(a) Normalise the ratings for this user. That is, compute the average rating and subtract it from individual ratings.

(b) Compute a user profile for the user, with components for processor speed, disk size, and main memory size.

Tutorial sheet Q5

(content-based method)

A certain user has rated the three computers as follows: A: 4 stars, B: 2 stars, C: 5 stars.

(a) The average rating is $(4+2+5)/3 = 11/3$. Therefore, the mean-centred ratings for each item are:

$A:\ 4 - 11/3 = 1/3$

$B:\ 2 - 11/3 = -5/3$

$C:\ 5 - 11/3 = 4/3$

Tutorial sheet Q5

Feature \ item	A	B	C
Processor speed	3.06	2.68	2.92
Disk size	500	320	640
Main-memory size	6	4	6

$A:\ 1/3, B: = -5/3, C:\ 5 - 11/3 = 4/3$

(b) weighted average of the component's values

Processor speed: feature × rating from user

$$3.06\times1/3 - 2.68\times5/3 + 2.92\times4/3 = 0.4467$$

Disk size:

$$500\times1/3 - 320\times5/3 + 640\times4/3 = 486.6667$$

Main memory size:

$$6\times1/3 - 4\times5/3 + 6\times4/3 = 3.3333$$

user profile

A	B	C	Feature 1	2	3
$\frac{1}{3}$	$-\frac{5}{3}$	$\frac{4}{3}$			