$\frac{9\frac{1}{2}}{10}$

Quiz #8: Recommendation Systems

Name: Yijun Lin ID: 3689281438

___Morning Session or ___Afternoon Session

1) (3pt) Using the user-based CF and considering all users, calculate the predicted rating for U_3 on item I_2 (using the Pearson correlation) (1pt). Do you use U_1 in your calculation and why (1pt)? What is the Jaccard similarity of user 1 and 3? (1pt)

3

$$w_{a,v} = \frac{\sum_{i \in I}(r_{a,i} - \bar{r}_a)(r_{v,i} - \bar{r}_v)}{\sqrt{\sum_{i \in I}(r_{a,i} - \bar{r}_a)^2}\sqrt{\sum_{i \in I}(r_{v,i} - \bar{r}_v)^2}} \qquad P_{a,i} = \bar{r}_a + \frac{\sum_{u \in U}(r_{u,i} - \bar{r}_u) \cdot w_{a,u}}{\sum_{u \in U} |w_{a,u}|}$$

	I_1	I_2	I_3	I_4
U_1	4	?	5	5
U_2	4	2	1	
U_3	3	?	2	4
U_4	4	4		
U_5	2	1	3	5

① $W_{2,3} = \frac{(4-\frac{5}{2})(3-\frac{5}{2}) + (1-\frac{5}{2})(2-\frac{5}{2})}{\sqrt{(4-\frac{5}{2})^2 + (1-\frac{5}{2})^2} \times \sqrt{(3-\frac{5}{2})^2 + (2-\frac{5}{2})^2}} = 1$

$W_{3,4} = \frac{(3-3)(4-4)}{\sqrt{(3-3)^2} \cdot \sqrt{(4-4)^2}} = 0$

$W_{3,5} = \frac{(3-3)(2-\frac{10}{3}) + (2-3)(3-\frac{10}{3}) + (4-3)(5-\frac{10}{3})}{\sqrt{(3-3)^2 + (2-3)^2 + (4-3)^2} \cdot \sqrt{(2-\frac{10}{3})^2 + (3-\frac{10}{3})^2 + (5-\frac{10}{3})^2}} = 0.65$

$P_{3,2} = \frac{9}{3} + \frac{(2-\frac{5}{2}) \times W_{2,3} + (1-\frac{10}{3}) \times W_{3,5} + (4-4) \times W_{3,4}}{|W_{2,3}| + |W_{3,4}| + |W_{3,5}|} = 1.78$

② I don't use U_1, because U_1 has no rating on I_2

③ Jaccard $(U_1, U_3) = 1$

2) (2pts) Using item-based CF (N=2) and the Pearson correlation, calculate the rating prediction of item 3 for user 1 using (1) average ratings based on all ratings (1pt) and (2) average ratings for co-rated items (1pt)

$1\frac{1}{2}$

Since N=2 you have to use only 2 values while calculating prediction.

$$w_{i,j} = \frac{\sum_{u \in U}(r_{u,i} - \bar{r}_i)(r_{u,j} - \bar{r}_j)}{\sqrt{\sum_{u \in U}(r_{u,i} - \bar{r}_i)^2}\sqrt{\sum_{u \in U}(r_{u,j} - \bar{r}_j)^2}} \qquad P_{u,i} = \frac{\sum_{n \in N} r_{u,n} w_{i,n}}{\sum_{n \in N} |w_{i,n}|}$$

	I1	I2	I3	I4
U1	2	1	~~?~~	3
U2	3	?	5	2
U3		4	2	3
U4	5	3	1	
	$\frac{10}{3}$	$\frac{8}{3}$	$\frac{8}{3}$	$\frac{8}{3}$

① $W_{1,3} = \frac{(3-\frac{10}{3})(5-\frac{8}{3}) + (5-\frac{10}{3})(1-\frac{8}{3})}{\sqrt{(3-\frac{10}{3})^2 + (5-\frac{10}{3})^2} \times \sqrt{(5-\frac{8}{3})^2 + (1-\frac{8}{3})^2}} = -0.73$

$W_{2,3} = \frac{(4-\frac{8}{3})(2-\frac{8}{3}) + (3-\frac{8}{3})(1-\frac{8}{3})}{\sqrt{(4-\frac{8}{3})^2 + (3-\frac{8}{3})^2}\sqrt{(2-\frac{8}{3})^2 + (1-\frac{8}{3})^2}} = 0.58$

$W_{3,4} = \frac{(5-\frac{8}{3})(2-\frac{8}{3}) + (2-\frac{8}{3})(3-\frac{8}{3})}{\sqrt{(5-\frac{8}{3})^2 + (2-\frac{8}{3})^2}\sqrt{(2-\frac{8}{3})^2 + (3-\frac{8}{3})^2}}$

$P_{1,3} = \frac{2 \times W_{1,3} + 1 \times W_{2,3} + 3 \times W_{3,4}}{|W_{1,3}| + |W_{2,3}| + |W_{3,4}|}$

② $W_{1,3} = \frac{(3-4)(5-3) + (5-4)(1-3)}{\sqrt{(3-4)^2 + (5-4)^2} \times \sqrt{(5-3)^2 + (1-3)^2}}$

$W_{2,3} = \frac{(4-\frac{7}{2})(2-\frac{3}{2}) + (3-\frac{7}{2})(1-\frac{3}{2})}{\sqrt{(4-\frac{7}{2})^2 + (3-\frac{7}{2})^2}\sqrt{(2-\frac{3}{2})^2 + (1-\frac{3}{2})^2}}$

$W_{3,4} = \frac{(5-\frac{7}{2})(2-\frac{5}{2}) + (2-\frac{7}{2})(3-\frac{5}{2})}{\sqrt{(5-\frac{7}{2})^2 + (2-\frac{7}{2})^2}\sqrt{(2-\frac{5}{2})^2 + (3-\frac{5}{2})^2}}$

$P_{1,3} = \frac{2 \times W_{1,3} + 1 \times W_{2,3} + 3 \times W_{3,4}}{|W_{1,3}| + |W_{2,3}| + |W_{3,4}|}$

→ use highest two weights since N = 2

3) (3pt) Briefly describe how Feature Augmentation (1pt), Feature Combination (1pt), and Meta-Level (1pt) Hybrid Recommendation Systems work

Feature Augmentation: Generate a new feature by a contributing recommender as augment profile. Pass the augment profile to another recommender

3

Meta-level: Use a model learned by one recommender as input for another

Feature Combination: Borrow recommedation logic from a virtual contributing recommender. inject the feature into the actual recommender

4) (2pts) Briefly explain one advantage (1pt) and one disadvantage (1pt) of using the dimensionality reduction techniques in your CF recommendation systems.

pro: the calculation would be ~~and~~ easier

2

con: some useful information may be lost