Tutorial Business Analytics

Homework 11

Exercise 11.3 - Collaborative Filtering

The following ratings of the users 1, 2, 3, and the active user 4 are given for the products A to J. The ratings can assume values in {1, 2, ..., 10}.

Product	User			Active User
	1	2	3	4
A	5	7	9	10
B	x	9	7	10
C	1	10	2	x
D	8	8	9	9
E	x	x	5	1
F	7	x	x	3
G	9	9	8	x
H	x	7	8	6
I	10	2	8	1
J	1	4	6	6

x: no rating available

$c_{a,i} = \frac{cov(r_a, r_i)}{\sigma_a \sigma_i}$

$c_{1,4} = \frac{-29.6}{12.37 \times 12.35} = -0.1984$

$w_{1,4} = s_{1,4} c_{1,4}$

$m = 5$

$s_{1,4} = 5/7$

$w_{1,4} = \frac{5}{7} \times -0.194 = -0.139$

For the significance weight of user a and i, $s_{a,i}$, and the number of co-rated items, m, we set:

$s_{a,i} = 1$, if $m \geq 7$

$s_{a,i} = \frac{m}{7}$, else

a) Determine the **weighted correlations**, $w_{i,4}$, between the users $i \in \{1,2,3\}$ and the active user $a = 4$.

b) Determine the **rating predictions**, $p_a(\text{product})$, for product C and product G, taking into account the two users most similar to the active user.

c) What problems are associated with the application of collaborative filtering in practice? Name at least three. Cold Start, Sparsity

d) What alternative recommending technique is there where these problems do not occur?

prediction formula

$p_{a,i} = \bar{r}_a + \frac{\sum_{u=1}^{k} r_{a,u}(r_{u,i} - \bar{r}_u)}{\sum_{u=1}^{k} |r_{a,u}|}$

$w_{2,4} = 0.7687$

$w_{3,4} = 0.5261$

$= 5.51$

$p_{4,C} = 4.6 + \frac{0.7687 \times (10 - 5.6)}{0.7687 + 0.5261} + \frac{0.5261 \times (2 - 6.2)}{0.7687 + 0.5261}$