

DS300.N11 - Hệ Khuyến Nghị
Bài tập học cộng tác dựa trên người dùng

Nguyễn Văn Kiệt, Huỳnh Văn Tín
Sinh viên: Phạm Đức Thế - 19522253

Thứ 3, ngày 13 tháng 09 năm 2022

Bài tập

	I_1	I_2	I_3	I_4	I_5
U_1	3	0	2	3	3
U_2	4	3	4	3	5
U_3	3	3	0	5	4
U_4	1	5	5	0	1
U_5	5	3	4	4	?

Tính giá trị rating của U_5 đối với I_5 ?

Chú ý: Dùng Pearson, tổng hợp đánh giá có trọng số, neighborhood size = 3

Lời giải.

- Tính rating trung bình:

$$\begin{aligned}\bar{r}_{U_1} &= \frac{3 + 0 + 2 + 3 + 3}{5} = 2.2 \\ \bar{r}_{U_2} &= \frac{4 + 3 + 4 + 3 + 5}{5} = 3.8 \\ \bar{r}_{U_3} &= \frac{3 + 3 + 0 + 5 + 4}{5} = 3.0 \\ \bar{r}_{U_4} &= \frac{1 + 5 + 5 + 0 + 1}{5} = 2.4 \\ \bar{r}_{U_5} &= \frac{5 + 3 + 4 + 4}{4} = 4.0\end{aligned}$$

- Tính độ tương đồng giữa 2 người dùng Pearson:

$$\begin{aligned}\text{sim}(U_5, U_1) &= \frac{(5-4) \times (3-2.2) + (3-4) \times (0-2.2) + (4-4) \times (2-2.2) + (4-4) \times (3-2.2)}{\sqrt{(5-4)^2 + (3-4)^2 + (4-4)^2 + (4-4)^2} \sqrt{(3-2.2)^2 + (0-2.2)^2 + (2-2.2)^2 + (3-2.2)^2}} \\ &= \frac{3}{\sqrt{2} \sqrt{\frac{154}{25}}} = 0.85 \\ \text{sim}(U_5, U_2) &= \frac{(5-4) \times (4-3.8) + (3-4) \times (3-3.8) + (4-4) \times (4-3.8) + (4-4) \times (3-3.8)}{\sqrt{(5-4)^2 + (3-4)^2 + (4-4)^2 + (4-4)^2} \sqrt{(4-3.8)^2 + (3-3.8)^2 + (4-3.8)^2 + (3-3.8)^2}} \\ &= \frac{1}{\sqrt{2} \sqrt{\frac{34}{25}}} = 0.61 \\ \text{sim}(U_5, U_3) &= \frac{(5-4) \times (3-3) + (3-4) \times (3-3) + (4-4) \times (0-3) + (4-4) \times (5-3)}{\sqrt{(5-4)^2 + (3-4)^2 + (4-4)^2 + (4-4)^2} \sqrt{(3-3)^2 + (3-3)^2 + (0-3)^2 + (5-3)^2}} \\ &= \frac{0}{\sqrt{2} \sqrt{13}} = 0\end{aligned}$$

$$\begin{aligned}
sim(U_5, U_4) &= \frac{(5-4) \times (1-2.4) + (3-4) \times (5-2.4) + (4-4) \times (5-2.4) + (4-4) \times (0-2.4)}{\sqrt{(5-4)^2 + (3-4)^2 + (4-4)^2 + (4-4)^2} \sqrt{(1-2.4)^2 + (5-2.4)^2 + (5-2.4)^2 + (0-2.4)^2}} \\
&= \frac{-4}{\sqrt{2} \sqrt{\frac{531}{25}}} = -0.61
\end{aligned}$$

- Tính rating của U_5 đối với I_5 bằng Tổng hợp đánh giá dựa trên khoảng cách đánh giá, giả sử neighborhood size = 3.

$$\begin{aligned}
r_{U_5, I_5} &= \bar{r}_{U_5} + \frac{sim(U_5, U_1)(r_{U_1, I_5} - \bar{r}_{U_1}) + sim(U_5, U_2)(r_{U_2, I_5} - \bar{r}_{U_2}) + sim(U_5, U_3)(r_{U_3, I_5} - \bar{r}_{U_3})}{sim(U_5, U_1) + sim(U_5, U_2) + sim(U_5, U_3)} \\
&= 4.0 + \frac{0.85 \times (3 - 2.2) + 0.61 \times (5 - 3.8) + 0 \times (4 - 3.0)}{0.85 + 0.61 + 0} = 4.97
\end{aligned}$$