

⑧ a) Similarity Jaccard = $\frac{|r_i \cap r_j|}{|r_i \cup r_j|}$ → تعداد فیلم‌هایی که دو user مشترکاً نمره داده‌اند،
 تعداد فیلم‌هایی که حداقل یکی از دو user نمره داده است.

Distance = 1 - similarity

$$\Rightarrow JD(u_1, u_2) = 1 - \frac{4}{8} = \frac{1}{2}, \quad JD(u_1, u_3) = 1 - \frac{4}{8} = \frac{1}{2}, \quad JD(u_2, u_3) = 1 - \frac{4}{8} = \frac{1}{2}$$

$$b) \underset{\text{cosine distance}}{CD}(v_i, v_j) = \arccos\left(\frac{v_i \cdot v_j}{\|v_i\| \cdot \|v_j\|}\right)$$

$$\Rightarrow CD(u_1, u_2) = \arccos\left(\frac{4}{\sqrt{8} \cdot \sqrt{8}}\right) = 48.24^\circ = CD(u_1, u_3) = CD(u_2, u_3)$$

c) اگر به جای blank صفر بنویسیم مشکل از قبیل مقادیر صفرها برای Jaccard داریم:

$$JD(u_1, u_2) = JD(u_1, u_3) = JD(u_2, u_3) = 0$$

$$d) CD(u_1, u_2) = \arccos\left(\frac{0+0+0+1+0+0+0+0}{\sqrt{4} \times \sqrt{2}}\right) = \arccos\left(\frac{\sqrt{2}}{4}\right) = 69.3^\circ$$

$$CD(u_1, u_3) = \arccos\left(\frac{0+0+0+1+0+0+1+0}{\sqrt{4} \times \sqrt{4}}\right) = \arccos\left(\frac{1}{2}\right) = 60^\circ$$

$$CD(u_2, u_3) = \arccos\left(\frac{0+0+0+1+0+0+0+0}{\sqrt{4} \times \sqrt{2}}\right) = 69.3^\circ$$

e)

	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8
u_1	$5 - \frac{19}{6}$	$4 - \frac{19}{6}$		$4 - \frac{19}{6}$	$1 - \frac{19}{6}$		$3 - \frac{19}{6}$	$2 - \frac{19}{6}$
u_2		$2 - \frac{13}{6}$	$4 - \frac{13}{6}$	$3 - \frac{13}{6}$	$1 - \frac{13}{6}$	$2 - \frac{13}{6}$	$1 - \frac{13}{6}$	
u_3	$2 - \frac{17}{6}$		$1 - \frac{17}{6}$	$3 - \frac{17}{6}$		$4 - \frac{17}{6}$	$4 - \frac{17}{6}$	$3 - \frac{17}{6}$

f)

	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8
u_1	$\frac{11}{6}$	$\frac{5}{6}$	$\frac{11}{6}$	$\frac{5}{6}$	$-\frac{7}{6}$	$-\frac{1}{6}$	$-\frac{7}{6}$	$\frac{1}{6}$
u_2		$-\frac{1}{6}$	$-\frac{11}{6}$	$\frac{1}{6}$		$\frac{7}{6}$	$\frac{7}{6}$	
u_3	$-\frac{5}{6}$							

$$CD(u_1, u_2) = \arccos \left(\frac{-\frac{5}{36} + \frac{25}{36} + \frac{91}{36} + \frac{7}{36}}{\sqrt{\frac{65}{8}} \times \sqrt{\frac{41}{8}}} \right) = 67.66^\circ$$

$$CD(u_2, u_3) = \arccos \left(\frac{-\frac{121}{36} + \frac{5}{36} - \frac{7}{36} - \frac{49}{36}}{\sqrt{\frac{41}{8}} \times \sqrt{\frac{41}{8}}} \right) = 133.63^\circ$$

$$CD(u_1, u_3) = \arccos \left(\frac{-\frac{55}{36} + \frac{5}{36} - \frac{7}{36} - \frac{7}{36}}{\sqrt{\frac{65}{8}} \times \sqrt{\frac{41}{8}}} \right) = 101.53^\circ$$

g)

	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8
u_1	1	1	0	1	0	0	1	0
u_2	0	0	1	1	0	0	0	0
u_3	0	0	0	1	0	1	1	1

\Rightarrow
 cluster 1 : m_1, m_2
 cluster 2 : m_3, m_5
 cluster 3 : m_4, m_7
 cluster 4 : m_6, m_8

h)

	C_1	C_2	C_3	C_4
u_1	4.5	1	3.5	2
u_2	2	2.5	2	2
u_3	2	1	3.5	3.5

i)

$$CD(u_1, u_2) = \arccos \left(\frac{9 + 2.5 + 7 + 4}{6.12 \times 4.27} \right) = 30.68^\circ$$

$$CD(u_2, u_3) = \arccos \left(\frac{4 + 2.5 + 7 + 7}{4.27 \times 5.43} \right) = 27.87^\circ$$

$$CD(u_1, u_3) = \arccos \left(\frac{9 + 1 + 12.25 + 7}{6.12 \times 5.43} \right) = 28.34^\circ$$