Recommender System
CS-4053

BASIL ALI KHAN
2016-0477

Part 1:

a)

	Item 1	Item 2	Item 3	Item 4	Item 5	Mean
user 1	?	3	?	3	4	3.33
user 2	4	?	?	2	?	3
user 3	?	?	3	?	?	3
user 4	3	?	4	?	3	3.33
user 5	4	3	?	4	4	3.75

$\Rightarrow$ User 1 & user 2

$$= \frac{(3-3.33)(2-3)}{\sqrt{(3-3.33)^2}\sqrt{(2-3)^2}} = 1$$

$\Rightarrow$ User 1 & user 4

$$= \frac{(4-3.33)(3.33)}{\sqrt{(4-3.33)^2}\sqrt{(3-3.33)^2}} = -1$$

$\Rightarrow$ User 1 & user 5

$$= \frac{(3-3.33)(3-3.75)+(3.33)(4-3.75)+(4-3.33)(4-3.75)}{\sqrt{(3-3.33)^2+(3-3.33)^2+(4-3.33)^2}\sqrt{(3-3.75)^2+(4-3.75)^2+(4-3.75)^2}}$$

$$= 0.49$$

Neighbour nearest to user 1 are user 2 & 5.

b) User 2 nearest neighbour.

user 1 & user2 = -1

used & user 4:

$$= \frac{(4-3)(3-3.33)}{\sqrt{(4-3)^2}\sqrt{(3-3.33)^2}} = -1$$

user 2 & user 5:

$$= \frac{(4-3)(4-3.75)+(2-3)(4-3.75)}{\sqrt{(4-3)^2+(2-3)^2}\sqrt{(4-3.75)^2+(4-3.75)^2}}$$

$$= 0$$

User 5 nearest neighbour.

User 1 & user 5 : 0.49

User 5 & user 2 : 0

User 5 & user 4:

$$= \frac{(3-3.33)(4-3.75)+(3-3.33)(4-3.75)}{\sqrt{(3-3.33)^2+(3-3.33)^2}\sqrt{(4-3.75)^2+(4-3.75)^2}}$$

$$= -1$$

$$R(U1, U3) = 33 + \frac{(4-3.33)(-1)}{|-1|}$$

$$= 2.66 \approx 3$$

	item 1	item 2	item 3	item 4	item 5
user 1	?	-0.33	?	-0.33	0.67
user 2	1	?	?	-1	?
user 3	?	?	0	?	?
user 4	-0.33	?	0.67	?	-0.33
user 5	0.25	-0.75	?	0.25	0.25

c) i) R (U3, I2)

item 2 & item 1 :

$$\text{Adj Cosine}(I2, I1) = \frac{(0.25)(-0.75)}{\sqrt{0.25^2}\ \sqrt{-0.75^2}} = -1$$

item 2 & item 3 :

Adj Cosine (I2, I3) = Not possible

item 2 & item 4 :

$$\text{Adj Cosine}(I2, I4) = \frac{(-0.33)(-0.33) + (-0.75)(0.25)}{\sqrt{-0.33^2 + (-0.75)^2}\ \sqrt{-0.33^2 + 0.25^2}}$$

$$= -0.23$$

item 2 & item 5 :

$$\text{Adj Cosine}(I2, I5) = \frac{(-0.33)(0.67) + (-0.75)(0.25)}{\sqrt{-0.33^2 + -0.75^2}\ \sqrt{0.67^2 + 0.25^2}}$$

$$= -0.69$$

item 4 & item 5 closest

ii) R(U3, I2) = Not possible

ii) R(U4, I4)

item 4 & item 1 : $\frac{(-1)(-1) + (0.25)(0.25)}{\sqrt{(-1)^2 + (0.25)^2}\sqrt{1^2 + 0.25^2}}$.

$= -0.90$

item 4 & item 2 : -0.23

item 4 & item 5 : $\frac{(0.33)(0.67) + (0.25)(0.25)}{\sqrt{(-0.33)^2 + (0.25)^2}\sqrt{(0.67)^2 + (0.25)^2}}$

$= -0.53$

Mean centered prediction function :

$R_{44} = 3.33 + \left\{ \frac{-0.28(-0.33)(-0.53)}{1-0.53} \right\}$

$= 3.66 \approx 4$

(d)

Using mode as a baseline depends on nature of structured of data. Model will give better and effective performance if data is highly skewed or has outliers. It can perform better in such scenario than mean.