1)

a)

Jaccard Similarity:

Sim(1,2) = 3/6 = 1/2

Sim(1,3) = 2/6 = 1/3

Sim(2,3) = 3/6 = 1/2

b)

Cosine similarity:

Sim(1,2)

= [(5\*3)+(5\*3)+(1\*1)]/[(42+52+52+12)1/2(32+42+32+12+22)1/2]

= 31/51.1175

= 0.606446

Sim(1,3)

= [(4\*2)+(5\*3)]/[ (42+52+52+12)1/2 (22+12+32+42)1/2]

= 23/44.8330

= 0.513015

Sim(2,3)

= [(4\*1)+(3\*3)+(2\*4)]/[(32+42+32+12+22)1/2(22+12+32+42)1/2]

= 21/34.2053

= 0.613941

c)

User 1 avg. = (4+5+5+1)/4 = 15/4

User 2 avg. = (3+4+3+1+2)/5 = 13/5

User 3 avg. = (2+1+3+4)/4 = 5/2

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Item 1 | Item 2 | Item 3 | Item 4 | Item 5 | Item 6 |
| User 1 | 1/4 | 5/4 |  | 5/4 | -11/4 |  |
| User 2 |  | 2/5 | 7/5 | 2/5 | -8/5 | -3/5 |
| User 3 | -1/2 |  | -3/2 | 1/2 |  | 3/2 |

d)

Centered-Cosine similarity:

Sim(1,2)

= 5.4/((3.2782)\*(2.28035))

= 5.4 / 7.47663

= 0.722251

Sim(1,3)

= 0.5/((3.2782)\*(2.23607))

= 0.5 / 7.33144

= 0.068199

Sim(2,3)

= -2.8/((2.28035)\*(2.23607))

= -2.8 / 5.09902

= -0.549125

Q2)

Summary of article: “Amazon.com Recommendations: Item-to-Item Collaborative Filtering”, by Greg Linden, Brent Smith, Jeremy York: