Analytical

1.

(A)

Jaccard Similarity problem;
Sim(1,2) = ½
Sim(1,3) = 1/3
Sim(2,3) = ½

(B)

Cosine Similarity Problem
Sim(1,2) = ((5*3) + (5*3) +(1*1)) / ((4²+5²+5²+1²)^{1/2} * (2²+1²+3²+4²)^{1/2}) → 0.606
Sim(1,3) = ((4*2) + (5*3)) / ((4²+5²+5²+1²)^{1/2} * (2²+1²+3²+4²)^{1/2}) → 0.513
Sim(2,3) = ((4*1)+(3*3)+(2*4)) / ((3²+4²+3²+1²+2²)^{1/2} * (2²+1²+3²+4²)^{1/2}) → 0.614

(C)

User 1 Avg	$(\frac{4+5+5+1}{4})$	15/4
User 2 Avg	$(\frac{3+4+3+1+2}{5})$	13/5
User 3 Avg	$(\frac{2+1+3+4}{4})$	5/2

User	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
1	¼	5/4		5/4	-11/4	
2		2/5	7/5	2/5	-8/5	-3/5
3	-½		-3/2	1/2		3/2

(D)

Centered-Cosine similarity:

Sim(1,2) = 5.4 / ((3.28) * (2.28)) → 0.72
Sim(1,3) = ½ / ((3.28)*(2.23607)) → 0.068
Sim(2,3) = -2.8 / ((2.28)*(2.23)) → -0.549

2. Report.

Zeid Al-Ameedi
03/01/2020
Cpts 315 HW2

<u>Two Decades of Recommender Systems at Amazon.com</u>

The following article centers around Amazon's recommendation system. This is the essence of users interacting and discovering related items that are either familiar or paired with the initially viewed product. Two decades in the making, the algorithm known as item-based collaborative filtering, launched in 1998. Since the success of the technique it has been adopted by major websites such as YouTube, Netflix and a variety of news sites.

The initial step of the algorithm is to identify and cross reference other users to find people with similar interests. These are usually associated similar purchase patterns which will identify related items. This is accomplished by iterating through each item in the catalog thus loosely defining "related" in said context. Pseudocode of the first step would look like:

For every item i^1:

   return I^2 = maxFrequency(i^1, item) # Essentially highest frequent item associated with item and i^1

The second step is a related items table being built off the returned information. Thus, the recommendation system is simplified to a simple lookup. Imagine all the possible combinations that can be generated basted on users current and previous behavior. The final step is to simply filter out the items already carted, seen or purchased. Maximizing revenue is easy when the user can be told what they want instead of relying them on hunting it down.