Morgan Baccus
CptS 315
Homework 2

Analytical Part

Q1.

a) Compute the Jaccard similarity between each pair of users:

Sim(1,2) = ½

Sim(1,3) = 1/3

Sim(2,3) = ½

b) Compute the cosine similarity between each pair of users:

Sim(1,2) = $[(5*3) + (5*3) + (1*1)] / [\sqrt{(4^2 + 5^2 + 5^2 + 1^2)} * \sqrt{(2^2 + 1^2 + 3^2 + 4^2)}] = 0.606$

Sim(1,3) = $[(4*2) + (5*3)] / [\sqrt{(4^2 + 5^2 + 5^2 + 1^2)} * \sqrt{(2^2 + 1^2 + 3^2 + 4^2)}] = 0.513$

Sim(2,3) = $[(4*1) + (3*3) + (2*4)] / [\sqrt{(3^2 + 4^2 + 3^2 + 1^2 + 2^2)} * \sqrt{(2^2 + 1^2 + 3^2 + 4^2)}] = 0.614$

c) Normalized matrix:

Avg(User 1) = (4 + 5 + 5 + 1) / 4 = 15/4

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

Avg(User 3) = (2 + 1 + 3 + 4) / 4 = 5/2

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

d) Compute the centered cosine similarity between each pair using the normalized matrix:

Sim(1,2) = 5.4 / (3.28 * 2.28) = 0.72

Sim(1,3) = -0.5 / (3.28 * 2.236) = 0.068

Sim(2,3) = -2.8 / (2.28 * 2.23) = -0.549

Q2.

This article gives an overview of the recommendation system that Amazon uses. Amazon has been developing their recommendation algorithm for over 20 years. The item-based collaborative filtering analyzes what items a user has viewed and suggests similar ones that that user might also be interested in. Other online-based companies have based their recommendation algorithm off of Amazon's since it is so successful.

The algorithm works by cross referencing the interests of a particular user with those of other users. The algorithm is able to assess which interests are similar and recommend the other users' items to the original user. The article goes on to describe how this is done in further detail. The next part of the algorithm is to create a table of the recommended items. The generated table allows the algorithm to sort by users' current and past behavior to quickly recommend items in the future. Once a user views something that is in this table, it is removed as to not be suggested again. When an online store such as Amazon has so many products, an algorithm like this can greatly increase the number of purchases made since it is easier and more efficient than a customer having to search for exactly what they want.