CS 345: Data Mining Names:_____

In-Class Activity 8 _____

Recommendation Systems

1. You want to design a recommendation system for an online bookstore that has been launched recently. The bookstore has over 1 million book titles, but its rating database has only 10,000 ratings. Which of the following would be a better recommendation system? a) User-user collaborative filtering b) Item-item collaborative filtering c) Content-based recommendation. Please justify your answer using at least one sentence.

2. Suppose you are using a content-based recommendation system. A customer has only rated two books: “Linear Algebra” and “Differential Equations” and both ratings are 5 out of 5 stars. Which of the following books is less likely to be recommended and why?
   a) “Operating Systems”
   b) “A Tale of Two Cities”
   c) “Convex Optimization”
   d) It depends on other users’ ratings.

3. Consider the following utility matrix.

	User 1	User 2	User 3	User 4	User 5
Item 1	4	4	4	1	1
Item 2	3	1		4	
Item 3	4	2		2	3
Item 4		2	3		1
Item 5			1	4	3
Item 6	1	1			2

a. Assume that we use the Pearson correlation coefficient as the similarity measure and that we predict a rating by averaging the two nearest (most similar) neighbors. Which two users do we use to predict the rating of Item 4 by User 1: _____

b. What is this predicted rating? _____

c. What is the correlation coefficient for each?_____