15.780, Fall 2021

Stochastic Models in Business Analytics

Problem Set 4 - Collaborative Filtering

Due date: November 19, 2021

Problem 1. (40 pts) Ted is gonna be DJ tonight, and he is really trying to impress one particular individual, let's call him User "6". He wants to play Rebecca Black's magnum opus ("Friday") for User 6 because he thinks they'll like it. Ted got a hold of the Spotify ratings for a few songs and people, including User 6. The data is shown in Table 1. Help him figure out if he should play "Friday" at the party.

Table 1: Song preferences

	Friday	Bad Blood	Work	Bohemian Rhapsody
User 1	-1	$+1$	-1	$+1$
User 2	-1	$+1$	-1	-1
User 3	$+1$	-1	$+1$	-1
User 4	-1	-1	-1	$+1$
User 5	$+1$	-1	$+1$	$+1$
User 6	???	-1	$+1$	-1

1. Calculate the cosine similarities for all users with User 6. (20 points)
2. Which two users are most similar to User 6? Does this make sense? Why or why not? (5 points)
3. Use a user based collaborative filter to find the predicted rating of User 6 for "Friday". You can threshold the user cosine similarities (from part 1) so that negative values become zero. (15 pts)