

CS432 Web Science: Assignment 6

Finished on April 2, 2020

Dr. Michele C. Weigle

Hannah Holloway
hhollowa@odu.edu

Contents

Problem 1	2
Question	2
Answer	2
Problem 2	3
Question	3
Answer	3
Problem 3	4
Question	4
Answer	4
Problem 4	4
Question	4
Answer	4

Listings

List of Figures

Problem 1

Question

Find 3 users who are closest to you in terms of age, gender, and occupation.

For each of those 3 users:

what are their top 3 (favorite) films? what are their bottom 3 (least favorite) films?

Based on the movie values in those 6 tables (3 users X (favorite + least favorite)), choose a user that you feel is most like you. Feel free to note any outliers (e.g., "I mostly identify with user 123, except I did not like "Ghost" at all"). You can investigate more than just the top 3 and bottom 3 movies to find your best match.

Answer

To get the 3 most liked and 3 most disliked films for three users that are the closest to me, I collected data from u.data, u.user and u.item. The combination of this data is shown below. I took this data and found users who are similar to me by being female, student, 21 . I first typed in programmer as the occupation but got 0 results. I changed my occupation to student, and was able to get 5 users. I picked the top three.

I used the loadMovieLens function provided in the reference code to get all the movies my top 3 users reviewed. I manually went through and picked the some of the highest and lowest rated for each user. As I ran each of my scripts, I appended the results to a text file.

My friend's most liked movies

Person's ID	Movie	Rating
198	'Full Metal Jacket (1987)'	10.680
198	'Godfather, The (1972)'	5.0
198	'2001: A Space Odyssey (1968)'	5.0
198	'Star Wars (1977)'	5.0
228	'Titanic (1997)'	5.0
228	'Good Will Hunting (1997)'	5.0
228	'English Patient, The (1996)'	5.0
324	'Nutty Professor, The (1996)'	5.0
324	'Air Force One (1997)'	5.0
324	'Toy Story (1995)'	5.0

My friend's least like movies

Person's ID	Movie	Rating
198	'Vegas Vacation (1997)'	1.0
198	'Vegas Vacation (1997)'	1.0
198	'Face/Off (1997)'	1.0
198	'Black Sheep (1996)'	1.0
228	'Cop Land (1997)'	1.0
228	'Big Night (1996)'	1.0
228	'Life Less Ordinary, A (1997)'	1.0
324	'Excess Baggage (1997)'	1.0
324	'Mad City (1997)'	3.0
324	'Mother (1996)'	3.0

After looking at all three user's likes and dislikes, I'd say the user that is the most like me is 198. It was a hard choice between 324 and 198, I could see that 324 gave movies like Homework Bound, The Nutty Professor, To Kill A Mockingbird and Toy Story which are some of my favorite movies. However, 198 had almost twice as more ratings as 324 which made the amount of shared dislikes and favorites higher. All of her top 5.0 choices I loved as well, she definitely my taste. All three of the users hated the same type of movies I did like Larry the Cable Guy, Vegas Vacation and Black Sheep. The third user, 228 had the fewest rating and a lot of the movie names I didn't recognize.

Problem 2

Question

Which 5 users are most correlated to the substitute you? Which 5 users are least correlated (i.e., negative correlation)?

Answer

By customizing the code "Making Recommendations" referenced I used the topMatched function and the similar person function to see the top 5 users similar to my substitute 198 as well as the top 5 least similar.

My results were: The 5 users that are most correlated to the substitute me: (1.0000000000000007, '926') (1.0, '88') (1.0, '86') (1.0, '818') (1.0, '762')

The 5 users that are least correlated to the substitute me: (-1.0, '611') (-1.0, '519') (-1.0, '36') (-1.0, '241') (-1.0, '191')

Problem 3

Question

Compute ratings for all the films that the substitute you has not seen.

Provide a list of the top 5 recommendations for films that the substitute you should see.

Provide a list of the bottom 5 recommendations (i.e., films the substitute you is almost certain to hate).

Answer

To get the top 5 recommendations, I used the getRecommendations portion of the code referenced.

The top 5 recommendations for user 198 are: (5.0, 'They Made Me a Criminal (1939)') (5.0, 'Star Kid (1997)') (5.0, 'Someone Else's America (1995)') (5.0, 'Santa with Muscles (1996)') (5.0, 'Saint of Fort Washington, The (1993)')

The bottom 5 recommendations for user 198 are: (1.0, 'August (1996)') (1.0, 'Amityville: Dollhouse (1996)') (1.0, 'Amityville: A New Generation (1993)') (1.0, 'Amityville 1992: It's About Time (1992)') (1.0, '3 Ninjas: High Noon At Mega Mountain (1998)')

Problem 4

Question

Choose your (the real you, not the substitute you) favorite and least favorite film from the data. For each film, generate a list of the top 5 most correlated and bottom 5 least correlated films.

Based on your knowledge of the resulting films, do you agree with the results? In other words, do you personally like / dislike the resulting films?

Answer

For this portion, I used the transformPrefs function on the reference code. Instead of comparing users, movies are compared. I was able to see the most similar and most different movies to my favorite movie on the list and least favorite. I chose Aristocats as my favorite and Cable Guy as my least.

My results were as follows:

The top 5 movies similar to "Aristocats, The (1970)" are: (1.0000000000000004, 'Fools Rush In (1997)') (1.0000000000000004, 'Associate, The (1996)') (1.0, 'Women, The (1939)') (1.0, 'Wings of the Dove, The (1997)') (1.0, 'Wild Bill (1995)')

The least 5 movies similar to "Aristocats, The (1970)" are: (-1.0, 'Deconstructing Harry (1997)') (-1.0, 'Calendar Girl (1993)') (-1.0, 'Blown Away (1994)') (-1.0000000000000007,

'When the Cats Away (Chacun cherche son chat) (1996)') (-1.0000000000000004, 'Before and After (1996)')

The top 5 movies similar to "Cable Guy, The (1996)" are: (1.00000000000000013, 'Bitter Moon (1992)') (1.0000000000000007, 'Heavy (1995)') (1.0, 'Wild Things (1998)') (1.0, 'Welcome To Sarajevo (1997)') (1.0, 'Wedding Gift, The (1994)')

The least 5 movies similar to "Cable Guy, The (1996)" are: (-1.0, "April Fool's Day (1986)") (-1.0, 'An Unforgettable Summer (1994)') (-1.0, 'Above the Rim (1994)') (-1.0, "'Til There Was You (1997)") (-1.0000000000000007, 'Flirt (1995)')

I don't recollect every seeing or hearing about any of these movies, but if they are similar to Cable Guy I'll stay away. I looked up the titles similar to Aristocats and was super confused. Not a single movie seemed on the list seemed remotely similar to the lighthearted children's movie. They were either depressing or crude. This might been caused by a fair outlier movie preference or maybe my substitute has similar taste and I the movies I didn't recognize her list were similar.

References