

Carnegie Mellon University in Qatar

AI for Medicine

15-182/282 - Spring 2022

Assignment 5

Name: _____

Andrew ID: _____

Due on: April 18, 2022 by midnight

Instructions:

- This assignment is only for 15-282 and it is optional (i.e., you do not need to submit it). It is meant only for you to practice on the included concepts.

Question	Points	Score
Recommending Doctors Using AI	50	
Total:	50	

Problem 1: Recommending Doctors Using AI (50 Points)

This problem is **ONLY** for 15-282. In it, you will build a recommendation system that recommends doctors to patients.

Table 1 captures the ratings (between 0 and 5, inclusive) of several patients for several doctors. For example, as shown in Table 1, Patient 1 has given a rating of 4.5 to Doctor 1, while Patient 4 has given a rating of 0 to the same doctor.

	Patient 1	Patient 2	Patient 3	Patient 4
Doctor 1	4.5	4.9	0	0
Doctor 2	3.8	4.6	0	0
Doctor 3	?	4.3	0	0
Doctor 4	0	0	4	5
Doctor 5	0	0	5	?

Table 1: Ratings of Patients for Doctors (ratings are between 0 and 5; ? entails no rating)

Answer the following questions, assuming a learning rate $\alpha = 0.5$, and make sure to show and submit all your work:

10pts

- (a) Suppose you are given the following table (i.e., Table 2) that captures certain values between 0 and 1 for two features x_1 and x_2 about each doctor in our dataset shown in Table 1.

	x_1	x_2
Doctor 1	1	0.1
Doctor 2	1	0
Doctor 3	0.8	0.2
Doctor 4	0.3	1
Doctor 5	0.1	1

Table 2

Predict the missing values in Table 1 (i.e., the ratings of Patient 1 for Doctor 3 and Patient 4 for Doctor 5) using a content-based recommendation algorithm. Write Python code to implement your algorithm. You can stop your algorithm after 10 rounds and make the predictions accordingly.

20pts

- (b) Suppose you are now given the parameter vectors for Patients 1, 2, 3, and 4 as follows:

$$\theta_1 = [0, 5, 0]$$

$$\theta_2 = [0, 5, 0]$$

$$\theta_3 = [0, 0, 5]$$

$$\theta_4 = [0, 0, 5]$$

On the flip side, assume you are NOT given the values of the two features x_1 and x_2 about each doctor (see Table 3).

Predict the missing values in Table 3 using a collaborative filtering algorithm. Write Python code to implement your algorithm. You can stop your algorithm after 15 rounds and make the predictions accordingly.

	x_1	x_2
Doctor 1	?	?
Doctor 2	?	?
Doctor 3	?	?
Doctor 4	?	?
Doctor 5	?	?

Table 3

For parts (c), (d), (e), (f), and (g) assume Patient 1 has given Doctor 3 a rating of 4.1 and Patient 4 has given Doctor 5 a rating of 5. Clearly, this completes Table 1.

- 4pts (c) Write Python code to produce and show the TF.IDF matrix that corresponds to Table 1.
- 4pts (d) Write Python code to produce and show the normalized TF.IDF matrix that corresponds to Table 1.
- 4pts (e) Who is the closest patient to Patient 4? Write Python code to answer this question (*Hint*: think about applying cosine similarity to rank all the patients against Patient 4).
- 4pts (f) Who is the closest doctor to Doctor 4? Write Python code to answer this question (*Hint*: think about applying cosine similarity to rank all the doctors against Doctor 4).
- 4pts (g) If the closest patient to Patient 4 has seen Doctor 4 before, but Doctor 4 is currently not available. Which doctor would you recommend for Patient 4 and why? Write Python code to answer this question.