Jeremy Kawahigashi

In-class Exercise 1

Collaborative Filtering by Hand

Companies that deliver streaming media, such as NETFLIX, Hulu, Qobuz, and Tidal, use machine learning to decide what to recommend to their customers based on the latter's usage records on their platform.

One of the techniques they use is collaborative filtering, which uses both the customer's own past data and data from other customers (who are, on average, somewhat likely to have seen movies or listened to songs that the customer in question might like).

Below is a small table of such data.

Table 1: Rating Table to Determine Recommendations for the Target User through Collaborative Filtering

RATING/ARTIST	Burna Boy	Chappell Roan	Laufey	Ne Obliviscaris	Nicki Minaj	Peso Pluma
Person 1	4	5	3	4	2	5
Person 2	3	4	3	2	4	4
Person 3	2	3	4	5	3	4
Person 4	3	4	4	5	2	5
Target User	4	?	3	?	4	2

How to do this exercise:

1. Form a group of two or three.

2. Take a quiet minute or two to look through the table separately.

3. When ready, share your guess as to the person most like the target user (based solely on this small dataset). Record all the guesses on a sheet of paper (one per person).

   Chappell Roan: 4

   Ne Obliviscraris: 3

4. If you all agree, explain (on your sheet) why you made this decision and whether you will or will not recommend Chappell Roan or Ne Obliviscaris to the target user. Include as much detail as you can about the reasoning you all used.

   Because the target user has similarities with person 1,2, match up on more than one artist, both were taken more heavily into account than person 3 and 4. In

regard for Chappell Roan, nobody on the table scored it lower than 3 and person 1,2 score it 5 and 4 respectively. Due to this Chappell Roan was scored as a 4.

Ne Obliviscraris was harder because person 2 rated it as 2. However person 3,4 both rated it as 5 and although they different slightly on music tastes, person 1 also rated it high with a 4. Taking all this into account I gave then a 3.

If you did not all agree, try to reach consensus by arguing for your position.

5. Once the discussion is done, record the reasons and arguments for each guess or match.

6. When you're done recording your process and the team's decision, think back to the first thing you did when you started solving the problem: What was the first thought or first piece of analysis? For example, did you start by reading the data for the target person or for the other four?

Read the target person and then try to find who matches most similarly

Did you use all four of the target person's entries or only one or two that stood out?
I used 2/4 for the most part

Did you look for matches vertically (down the columns) or horizontally (all ratings for one person, then the next person, etc.), or some other way.

Horizontally to see which artist were liked by which person then compare to the target

Once you started seeing a pattern, did you have doubts?

Yes

Were there data points that went against what you first saw? How did you deal with that?

Yes, and for those I took the general rating of the person while docking points because of the outlier

Also, think of more questions like the ones I asked above and explain the thought processes you went through.

After reviewing the data, is there anything else that could have affected your outcome?

Yes both Nicki and Peso both had varying scores when comparing person 1 and 2 with the target person.