Lesson 7: Collaborative Filtering — Reviewer

1 Introduction to Collaborative Filtering

1.1 Core Concept

The fundamental logic behind collaborative filtering can be described as follows:

- If an item is **Liked by Alice and Bob**, this implies they are ***Similar Users***.
- Consequently, if an item is **Liked by Alice**, it can be **Recommended to Bob**.
- This process is known as ***Collaborative learning***.

2 What is a Recommendation System?

There are a lot of applications where websites collect data from their users and use that data to ***predict the likes and dislikes*** of their users. This allows them to ***recommend the content that they like***. Recommender systems are a way of suggesting similar items and ideas to a user's specific way of thinking.

2.1 Two Types of Recommendation Systems

1. ***Collaborative Filtering***: Collaborative Filtering recommends items based on ***similarity measures*** between users and/or items.
2. ***Content-Based Recommendation***: It is ***supervised machine learning*** used to induce a classifier to discriminate between interesting and uninteresting items for the user.

3 What is Collaborative Filtering?

In Collaborative Filtering, we tend to ***find similar users*** and recommend what similar users like. In this type of recommendation system, we don't use the features of the item to recommend it; rather, we ***classify the users into clusters*** of similar types and recommend each user according to the preference of its cluster.

3.1 4 Types of Collaborative Filtering

- ***Memory-Based***
- ***Model-Based***
- ***Hybrid***
- ***Deep Learning***

3.2 Collaborative Filtering Visually

(Visual Flow Interpretation):

- Identify ***similar*** users.
- User A **buys** an item; User B **buys** a similar item.
- The system generates a ***recommend***ation (e.g., COLA).

4 Measuring Similarity in Collaborative Filtering

The following table illustrates how users rate different movies, which helps in measuring similarity:

Users	Movie 1	Movie 2	Movie 3	Movie 4
User 1	5	4		5
User 2	4		3	
User 3		1		2
User 4	1	2		

5 Rounding the Data

In collaborative filtering, we ***round off the data*** to compare it more easily. For example, we can assign ratings ***below 3 as 0*** and ratings ***above it as 1***. This will help us to compare data more easily.

Example of Rounded Data:

Users	Movie 1	Movie 2	Movie 3	Movie 4
User 1	1	1		1
User 2	1		1	
User 3		0		0
User 4	0	0		

6 Normalizing Rating

In the process of normalizing, we take the ***average rating of a user*** and ***subtract all the given ratings from it***. This gives us either positive or negative values as a rating, which can simply classify further into similar groups.

By normalizing the data, we can ***make clusters of the users*** that give a similar rating to similar items, and then we can use these clusters to recommend items to the users.

6.1 Step-by-Step Example

1. Compute the average rating for each user.

- If a user rated 3 movies as 4, 5, and 2, their average is:

$$(4 + 5 + 2)/3 = 3.67$$

2. Subtract the *average rating* from each rating:

- Movie 1: $4 - 3.67 = +0.33$
- Movie 2: $5 - 3.67 = +1.33$
- Movie 3: $2 - 3.67 = -1.67$

Now, instead of using the raw ratings, we use these ***normalized ratings***.

7 Reference

- https://www.geeksforgeeks.org/machine-learning/collaborative-filtering-ml/