SNS COLLEGE OF ENGINEERING

Kurumbapalayam (Po), Coimbatore – 641 107

An Autonomous Institution

Accredited by NAAC – UGC with 'A' Grade

Approved by AICTE, New Delhi & Affiliated to Anna University, Chennai

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

COURSE NAME : 19CS508 - BIG DATA ANALYTICS

III YEAR / V SEMESTER

Unit 3- ASSOCIATION AND RECOMMENDATION SYSTEM

Topic 4 : Collaborative Based Filtering or Recommend system

Collaborative Based Filtering or Recommend system

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.

There are basically four types of algorithms o say techniques to build Collaborative filtering recommender systems:

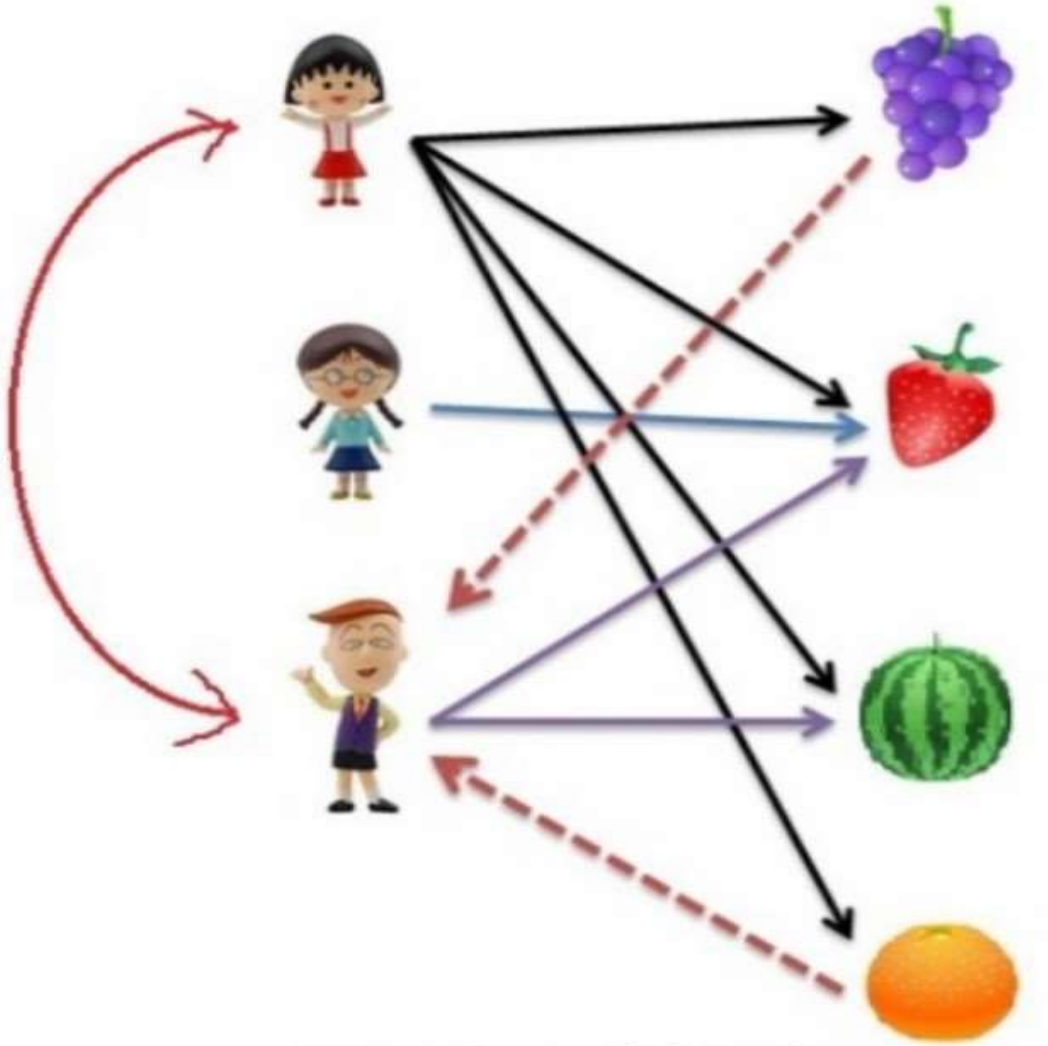
- Memory-Based
- Model-Based
- Hybrid
- Deep Learning

Cont..

Recommending the new items to users based on the interest and preference of other similar users is basically collaborative-based filtering. For eg:- When we shop on Amazon it recommends new products saying *"Customer who brought this also brought"* as shown below.

This approach not only addresses the limitations of content-based filtering but also leverages user interactions, making it more robust. By focusing on the historical performance of users, this recommendation system can predict future preferences with greater accuracy.

Cont..



Cont..

Measuring Similarity

A simple example of the movie recommendation system will help us in explaining:

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		

Cosine Similarity

We can also use the cosine similarity between the users to find out the users with similar interests, larger cosine implies that there is a smaller angle between two users, hence they have similar interests.

We can apply the cosine distance between two users in the utility matrix, and we can also give the zero value to all the unfilled columns to make calculation easy, if we get smaller cosine then there will be a larger distance between the users, and if the cosine is larger than we have a small angle between the users, and we can recommend them similar things.

$$Cosine\ (x,y) = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2}\ \sqrt{\sum_{i=1}^{n} y_i^2}}$$

Activity

Advantages of Collaborative Based Filtering or Recommend system

- No domain knowledge is required since all the features are learned automatically.
- Can help users discover new interests even if they're not actively searching for them by recommending new items similar to what they're interested in.
- Does not require in-detail features and contextual data of products or items. It only needs the user-item interaction matrix to train the matrix factorization model.

Disadvantages of Collaborative Based Filtering or Recommend system

- Data sparsity can lead to difficulty in recommending new products or users since the suggestions are based on historic data and interactions.
- As the user base grows, the algorithms suffer due to high data volume and lack of scalability.
- Lack of diversity in the long run. This might seem counterintuitive since the whole point of collaborative filtering is to recommend new items to the user. However, since the algorithms function based on historical ratings, it will not recommend items with little or limited data. Popular products will be more popular in the long run and there will be a lack of new and diverse options.
- .

Assessment 1

1. List out the advantages of Collaborative Based Filtering or Recommend system

a)_____

b)_____

c)_____

d)_____

2. Identify the disadvantages of Collaborative Based Filtering or Recommend system

a)_____

b)_____

c)_____

d)_____

REFERENCES

1. Anand Rajaraman and Jeffrey David Ullman, "Mining of Massive Datasets", Cambridge University Press, 2012.
2. David Loshin, "Big Data Analytics: From Strategic Planning to Enterprise Integration with Tools, Techniques, NoSQL, and Graph", Morgan Kaufmann/Elsevier Publishers, 2013
3. Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data", Wiley publishers, 2015.
4. https://www.ibm.com/topics/collaborative-filtering

THANK YOU