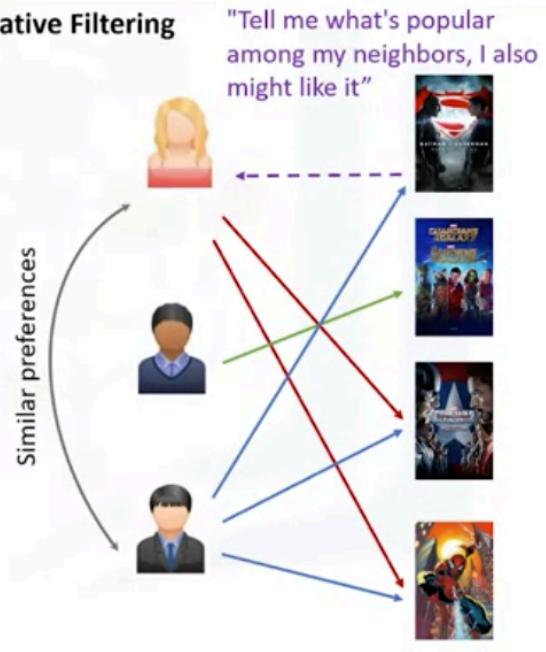


Two types of recommender systems

Content-Based



Collaborative Filtering



Implementing recommender systems

• Memory-based

- Uses the entire user-item dataset to generate a recommendation
- Uses statistical techniques to approximate users or items
e.g., Pearson Correlation, Cosine Similarity, Euclidean Distance, etc.

• Model-based

- Develops a model of users in an attempt to learn their preferences
- Models can be created using Machine Learning techniques like regression, clustering, classification, etc.

Content based recommendation system tries to recommend items to the users based on their profile.

Content-based recommender systems



Weighing the genres

Weighted Genre Matrix				
	Comedy	Adventure	Super Hero	Sci-Fi
	0	2	2	0
	10	10	10	10
	8	0	8	0

Input User Ratings	Movies Matrix				User Profile
	Comedy	Adventure	Super Hero	Sci-Fi	
	0	1	1	0	
	1	1	1	1	
	1	0	1	0	

	Comedy	Adventure	Super Hero	Sci-Fi
	18	12	20	5

COGNITIVE CLASS

Normalize the user profile:

User Profile	Comedy	Adventure	Super Hero	Sci-Fi
	0.3	0.2	0.33	0.16

Finding the recommendation



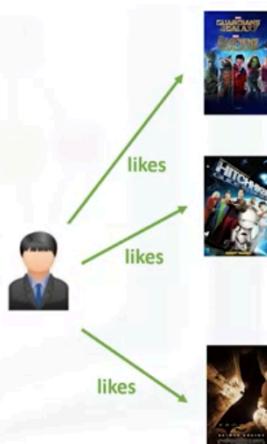
The recommendation in a content-based system is based on user's taste and the content or feature set items. Such a model is very efficient. However, in some cases, it doesn't work.

For example, assume that we have a movie in the drama genre, which the user has never watched. So, this genre would not be in her profile. Therefore, she will only get recommendations related to genres that are already in her profile and the recommender engine may never recommend any movie within other genres. This problem can be solved by other types of recommender systems such as collaborative filtering.

Collaborative filtering

- **User-based collaborative filtering**
 - Based on users' neighborhood

- **Item-based collaborative filtering**
 - Based on items' similarity



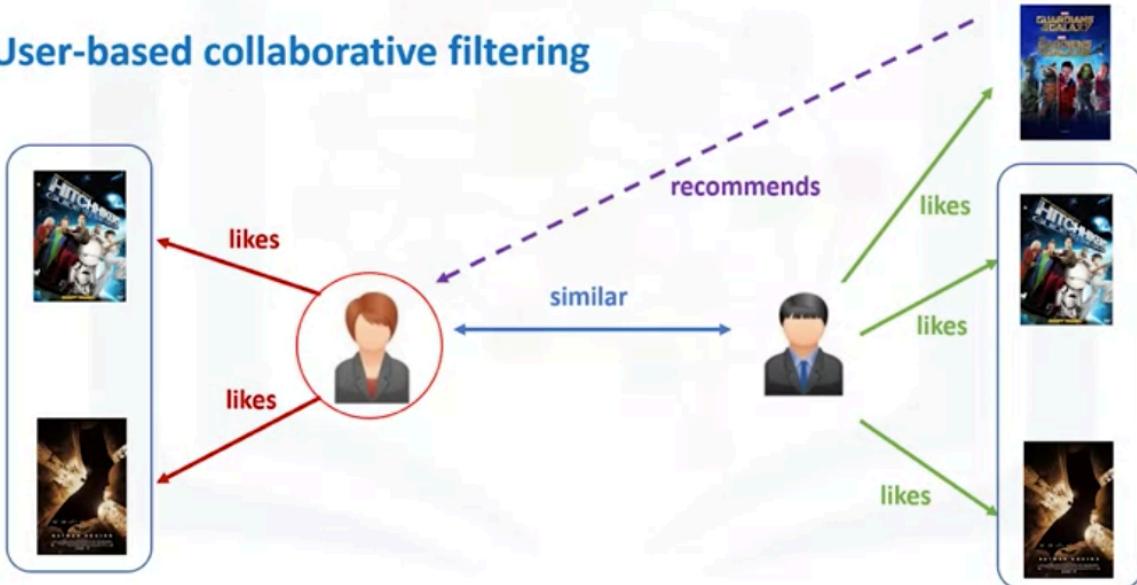
In user-based approach, the recommendation is based on users of the same neighborhood, with whom he/she shares common preferences.

Learning the similarity weights



Collaborative filtering

- User-based collaborative filtering**



These numbers represent similarity weights or proximity of the active user to other users in the

dataset.

Creating the weighted ratings matrix

The diagram illustrates the calculation of weighted ratings from three matrices:

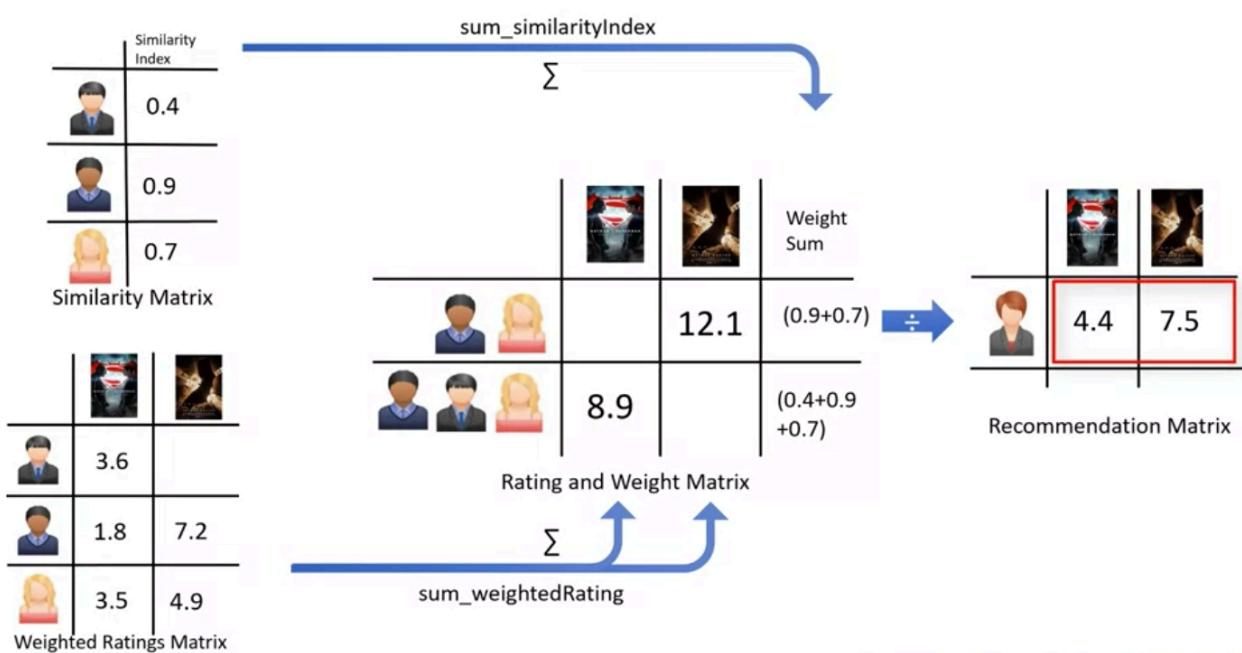
- Ratings Matrix Subset:**

	9	
	2	8
	5	7
- Similarity Matrix:**

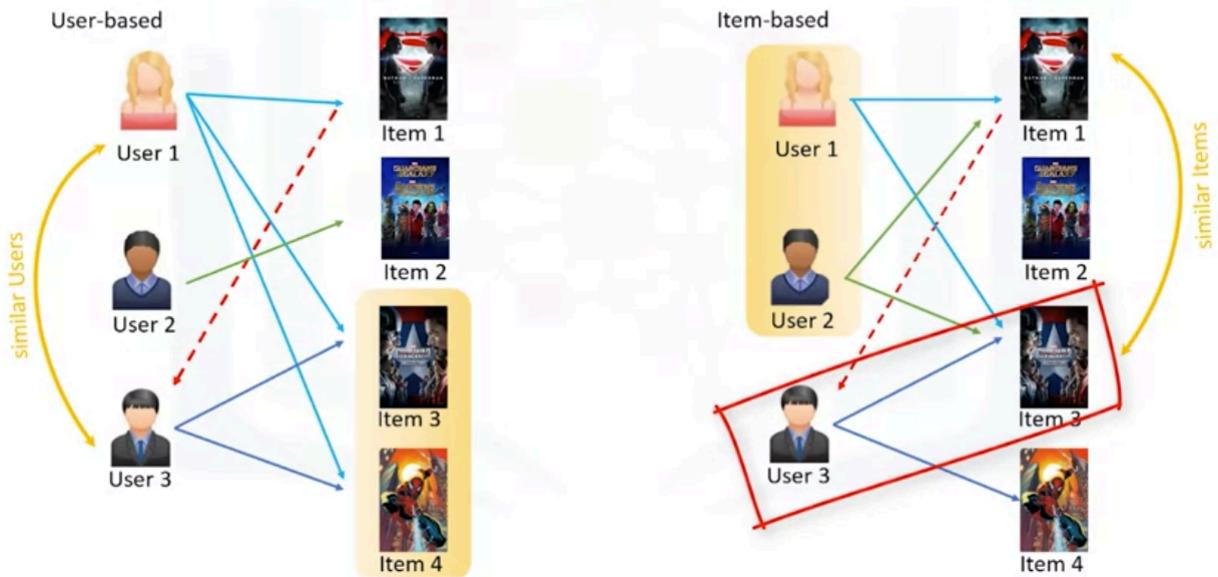
		Similarity Index
		0.4
		0.9
		0.7
- Weighted Ratings Matrix:**

	3.6	
	1.8	7.2
	3.5	4.9

The process involves multiplying each user's ratings by their similarity index and summing the results for each movie.



Collaborative filtering



In the item-based approach, similar items build neighborhoods on the behavior of users. Please note however, that it's not based on their contents. For example, item 1 and 3 are considered neighbors as they were positively rated by both User 1 and User 2. So Item 1 can be recommended to User 3 as he has already shown interest in Item 3. Therefore, the recommendations here are based on the items in the neighborhood that a user might prefer.

Challenges of collaborative filtering

- **Data Sparsity**
 - Users in general rate only a limited number of items
- **Cold start**
 - Difficulty in recommendation to new users or new items
- **Scalability**
 - Increase in number of users or items

As the number of users or items increases and the amount of data expands, collaborative filtering

algorithms will begin to suffer drops in performance, simply due to growth and the similarity computation. There are some solutions for each of these challenges such as using hybrid based recommender systems.