**RECOMMENDER SYSTEMS**

The recommender systems are considered as one of the most important tool for some business to expand their worth by providing personalizing experiences to the customers .

So what exactly recommender system does such that it becomes a buzz word for personalizing experience providing business. The recommender systems capture the pattern of peoples’ behavior and use it to predict what else they might want or like.

This systems are used to us in day to day life while purchasing anything through amazon , flipkart or watching movies on Netflix. On our previous history of movies watched and rating given by us , Netflix system recommends us some of the movies we might like to watch based on our tastes.

So we can see applications in many places .Some of the applications are as follows:-

-What to buy?

E-commerce ,books ,movies ,beer ,shoes etc.

-Where to eat?

-Which job to apply to?

-Personalize your experience on the web

News platforms, news personalization

By using this recommender system we get some advantages which are listed out as follows:-

-Broader exposure

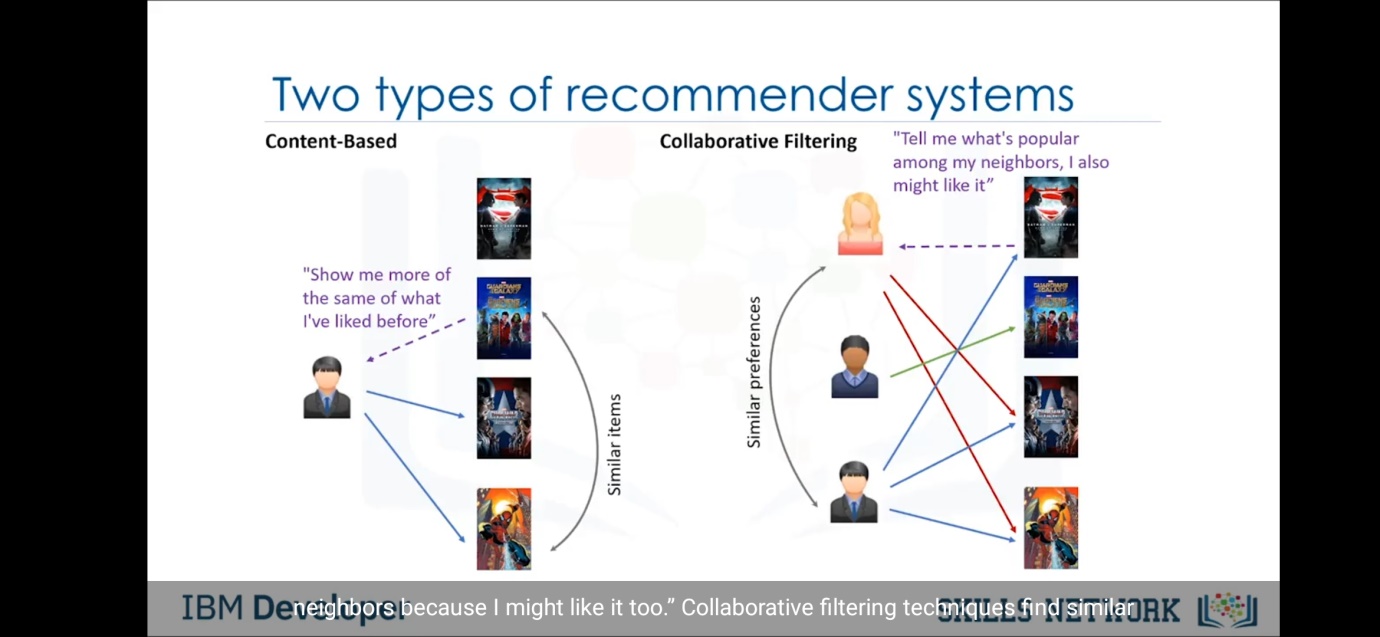
-Possibility of continual usage of products

-Provides better experience

**Types of recommender systems:-**

There are two types of recommender systems i.e. content based and collaborative based filtering.

Content based filtering recommends you the similar items you have used. Let’s take example we have list of movies form which you have watched 2 films. Now in content based filtering the systems recommends you the movie based on the 2 movies you have watched. So it’s a phrase used by user as ‘Show me more of the same of what I’ve liked before’.

On the other way , Collaborative filtering recommends you the film you might like based on the similar users liked. Continuing with the example , in this system the user will be recommended movies based on the similar users who like some movies that the actual user haven’t watched yet. Suppose there are 3 users a ,b and c. Now a has watched 2 films, b has watched 3 films out of which 2 are the films watched by user a. So this system will first find that both users liked the respective 2 films , so can been considered as similar users. Now the user b liked the third film he/she watched. SO that film will be recommended to user a , because the taste of user a and b matches or they are “Similar users”. 

The implementation of recommending systems can be done by two ways.

1) Memory based and 2) Model based

**Memory Based**

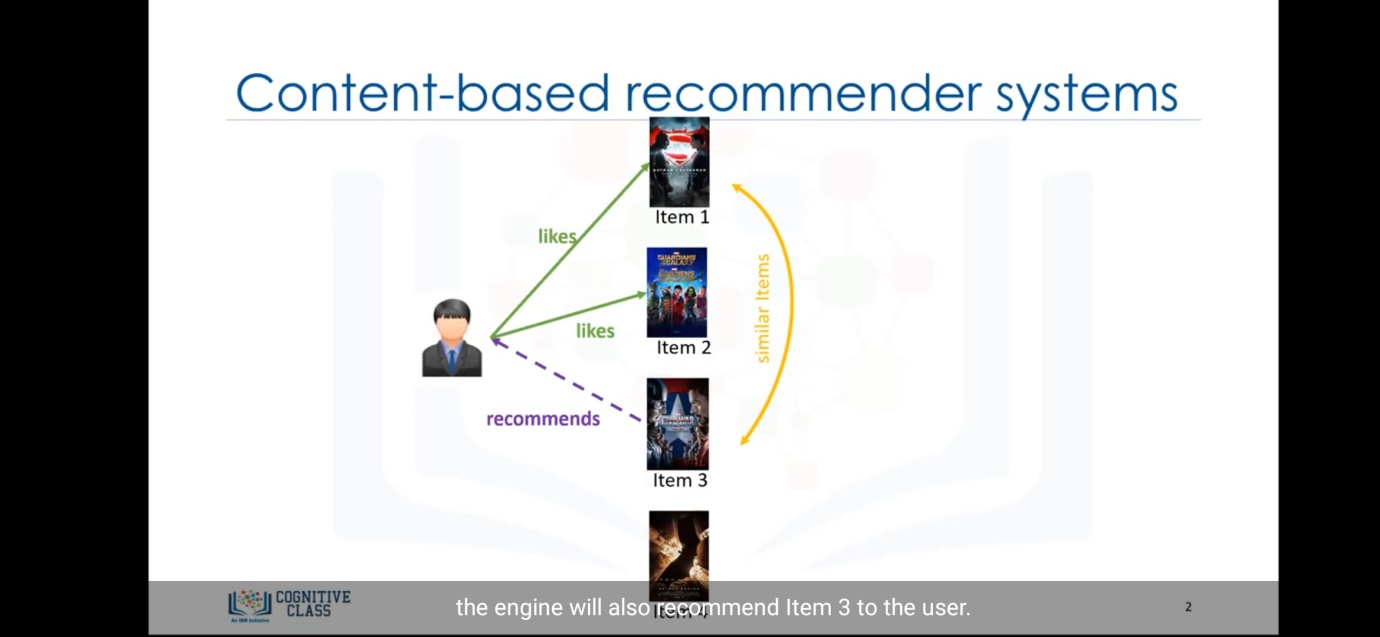
This method uses the entire user –item dataset to generate a recommendation .It uses statistical techniques to approximate users or items e.g. Pearson correlation ,Cosine similarity ,Euclidian distance etc.

**Model Based**

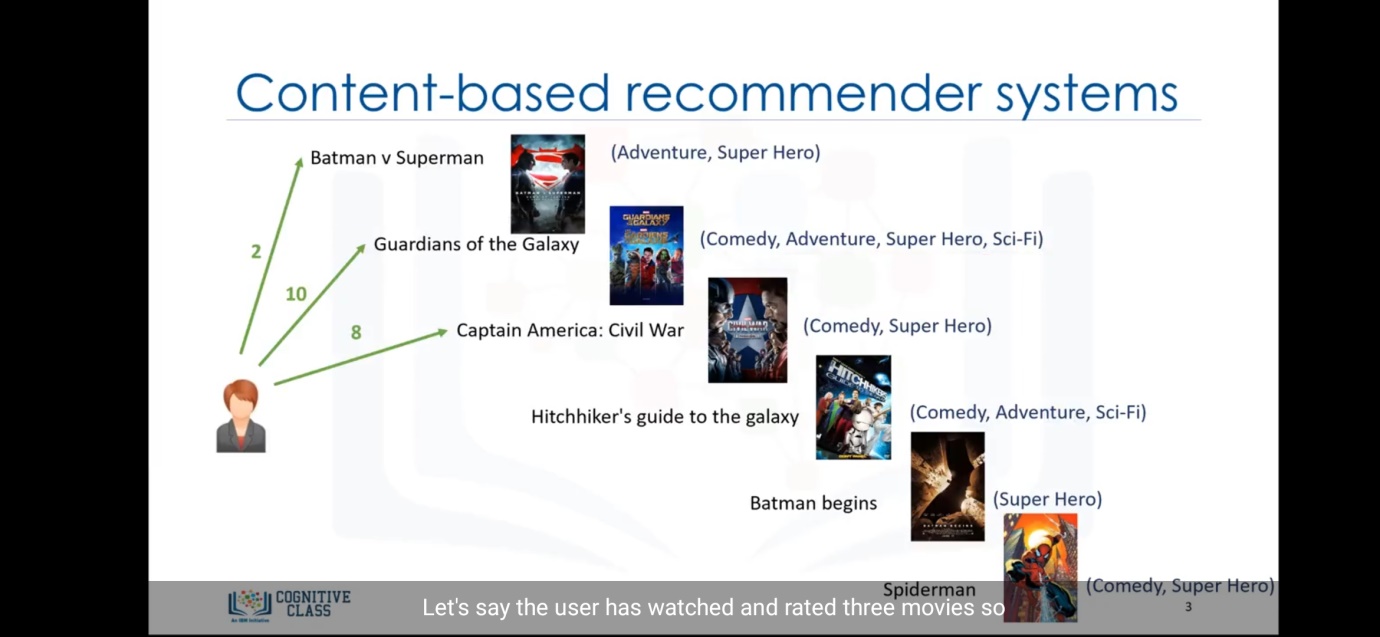
This method devlops a model of users in an attempt to learn their preferences. Models can be created using the Machine Learning techniques like regression,clustering,classification etc.

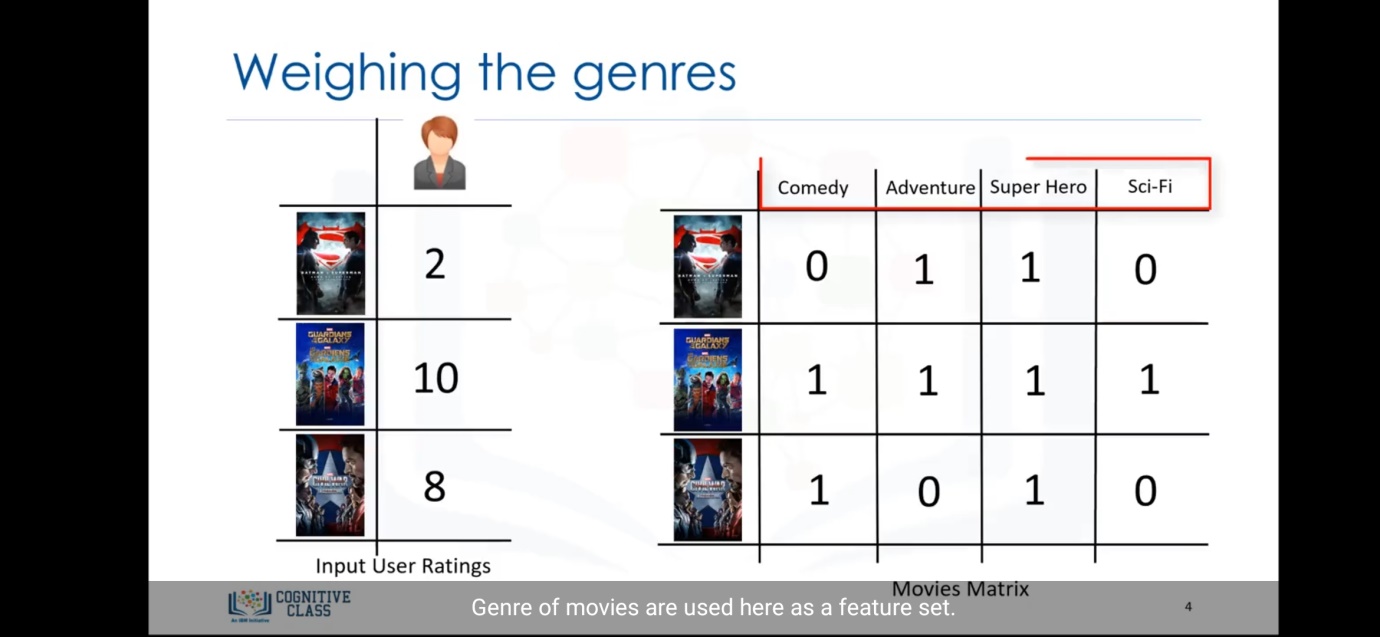
Now we shall study the first type of recommending systems i.e. content based recommender systems , in depth.

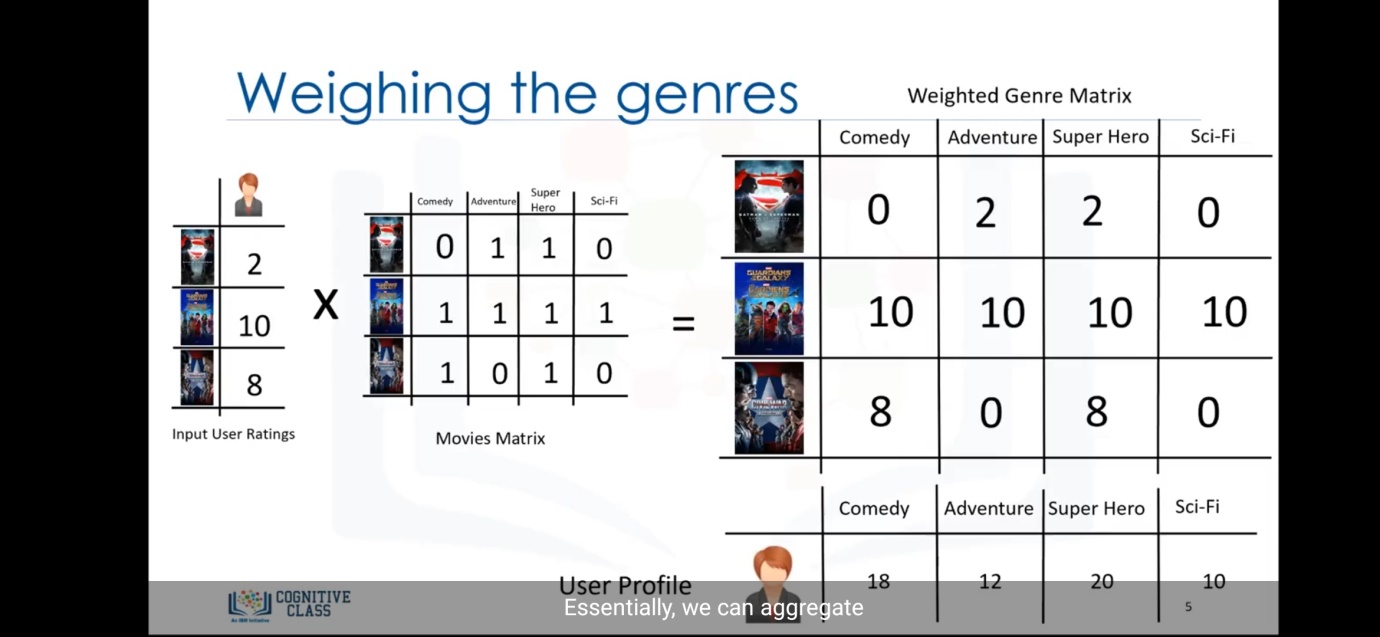
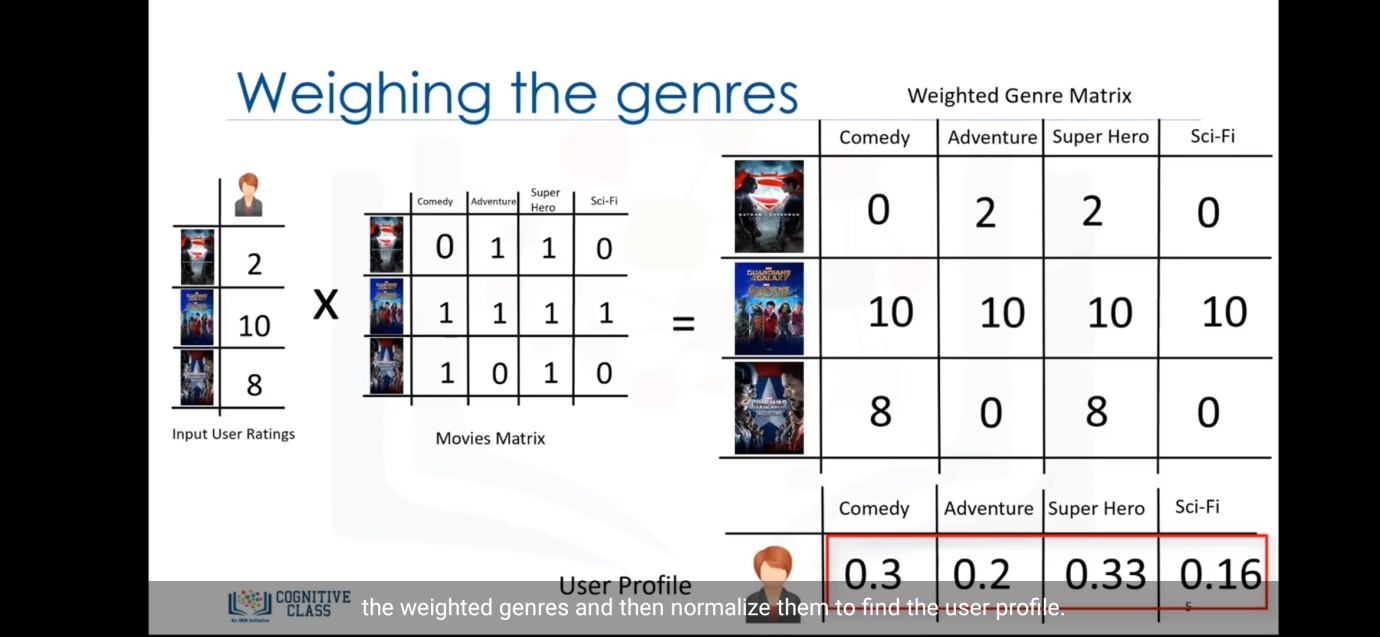
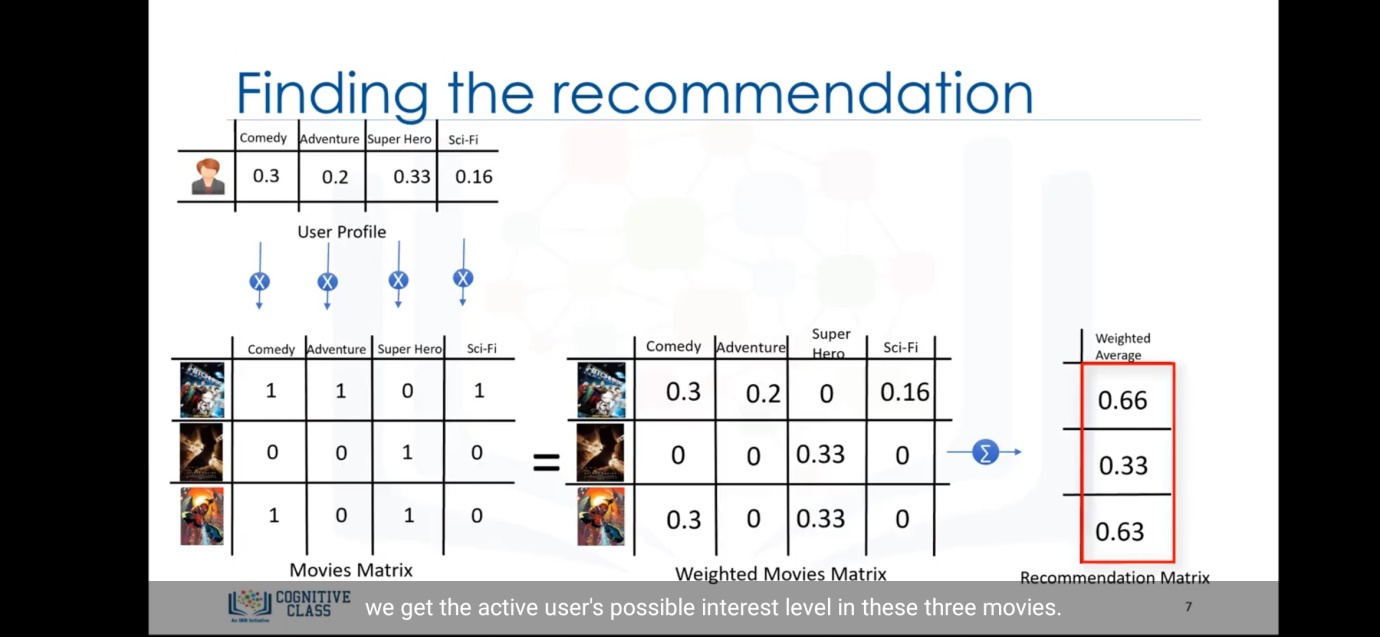
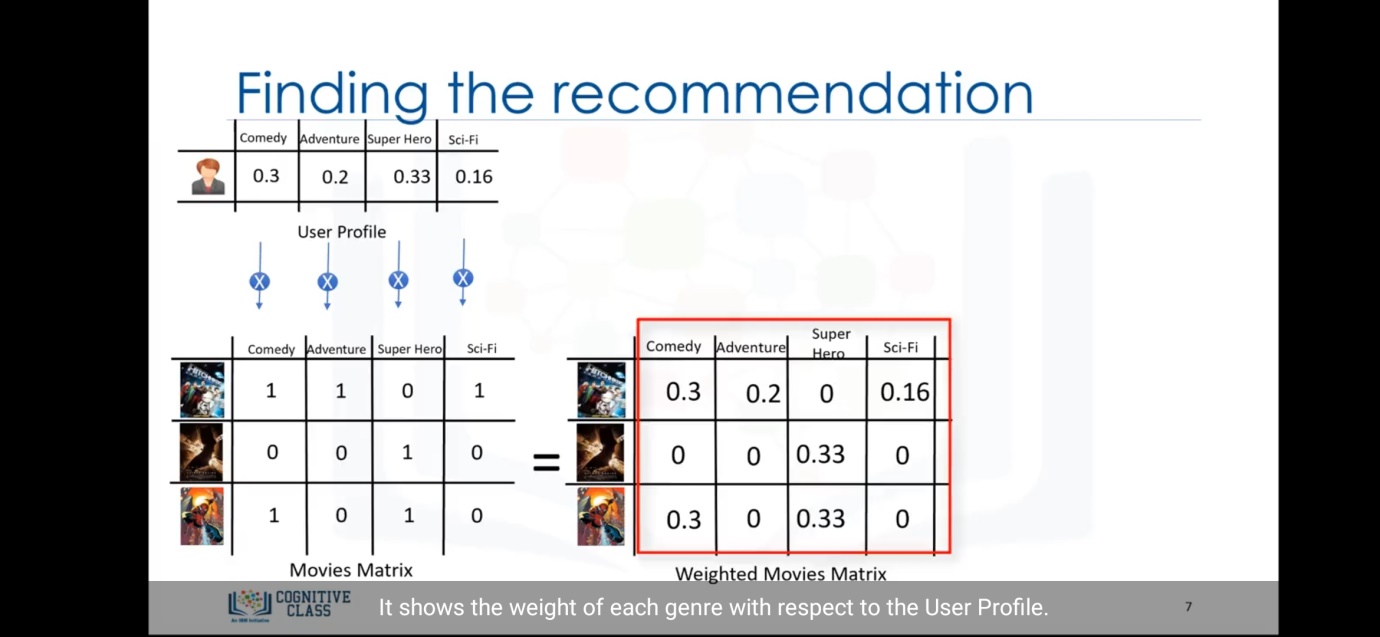
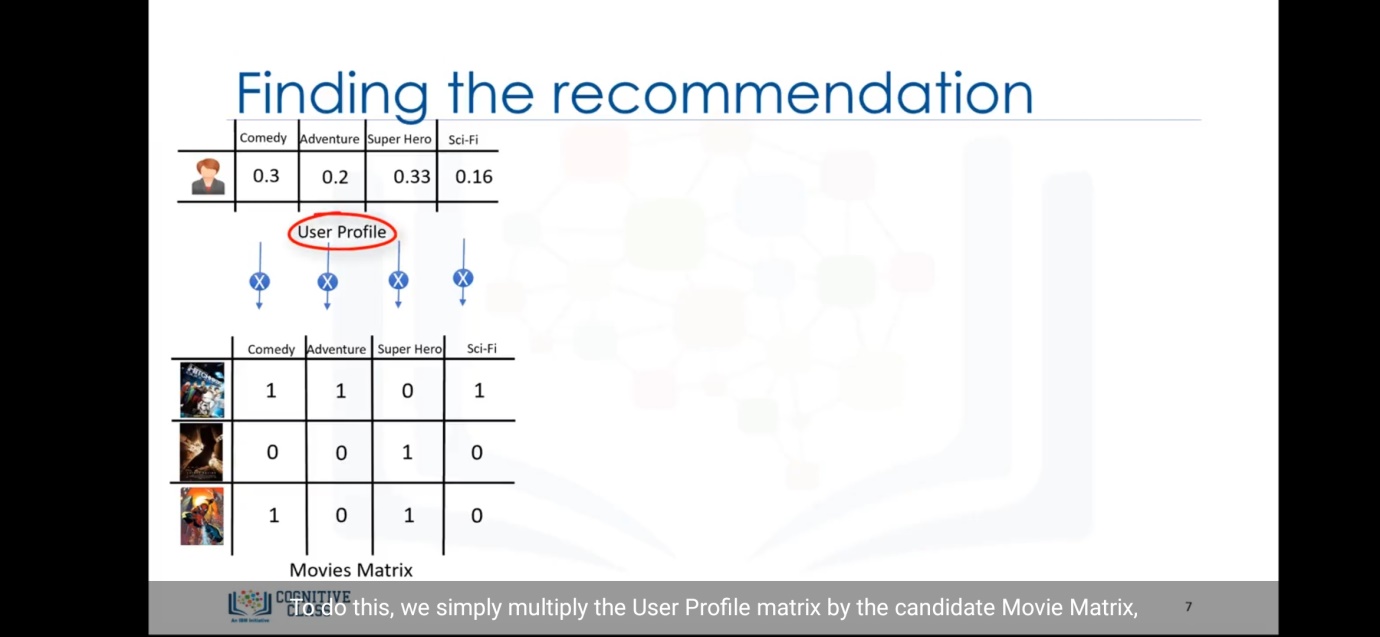
So general idea behind this recommendation system is that it recommends user the items similar to the previous ones that user liked.As shown in the figure , there are 4 films out of which user liked first and second films. Now the user haven’t watch the third and fourth films. So system will recommend third film as it is similar film based on genre user liked from previous first two films.



Now lets go through the algorithm.

Suppose there are 6 films as shown in the figure with their given respective genres. Now the user have watched first three films and given ratings according to their respective likes/dislikes . Here shown that user gives 3/10 for film “Batman v Superman”,10/10 for “Guardians of galaxy”,8/10 for “Captain America”. And the other films are yet to watch by user, so we have to recommend films from this leftover based on the ratings given by user to the 3 films. 

Now the first step to perform is that we have to generate a movies matrix which consist of the films user have watched and global genres.In our case genres are comedy , adventure , superhero and sci-fi. Now if the movie has that genre then value in matrix should be 1 else 0.So for (i,j) where i is movie and j is genre , matrix value of (i,j) is 1 if movie consist of that genre else value is 0.Refer the image for the matrix formed in this example.

Now we multiply the rating matrix with the movies matrix to get the weighted genre matrix.After getting the weighted genre matrix, add all the values with respect to the genre. So suppose comedy genre has 0 points for film “Batman v Superman” ,10 points for “Guardians of galaxy” and 8 point in “Captain america” , so value of comedy genre will be 0+10+8=18. After that we normalize the data and we get results shown in the figure.Hence we got the matrix which state that the user likes comedy with value 0.3 , adventure with 0.2,Superhero with 0.33 and scifi by 0.16.Let this matrix be called the user profile matrix.Now we will form the movie matrix for the movies the user haven’t watched yet. So we will further multiply the user matrix and movies genre matrix(movies yet to watch) and we get the weighted move matrix again. Whose sum we will calculate as sum of all the values of a row respected to that movie.Three images are shown below which shows all this results.First image calculates mavies matrix,second one finds the weighted matrix obtained by multipling user matrix and movies matrix. And the third one shows the weighted average obtaind by sum.

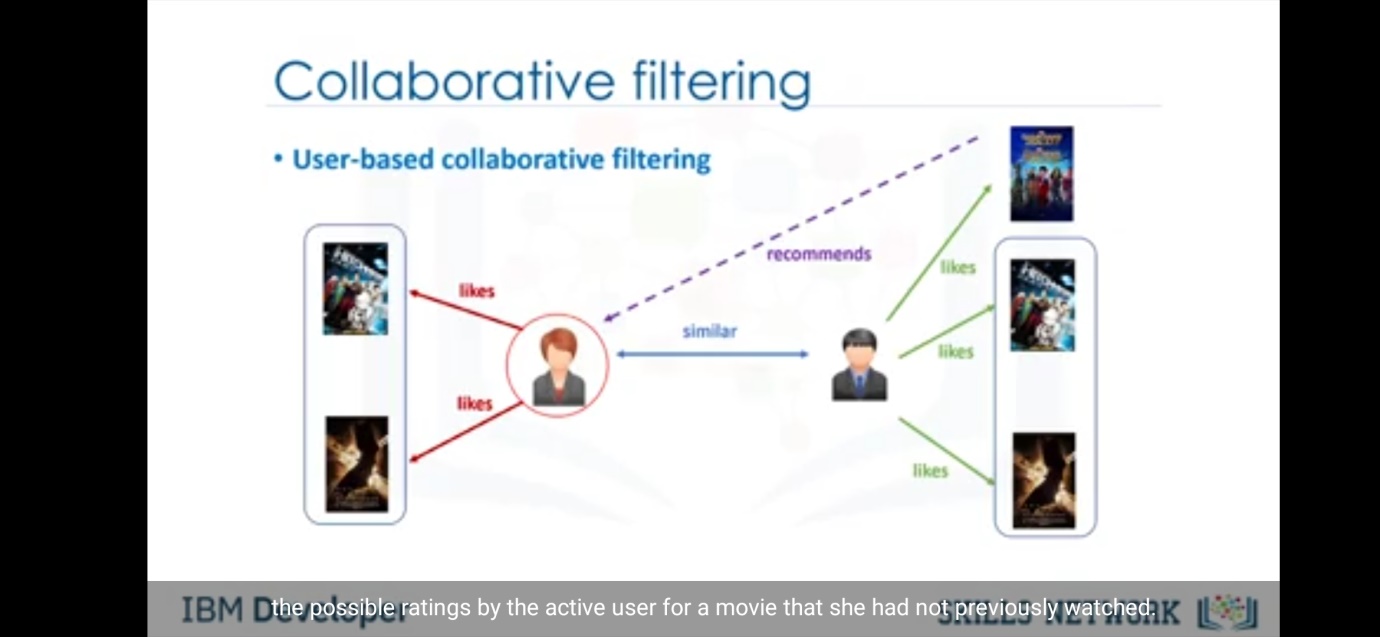
By this we got values 0.66,0.33 and 0.33 which are normalized and counted out of 10. So we get approx rating the user should gave after watching those movies as 6.6,3.3 and 3.3. So from this the one with the higher rating can be recommended easily. Hence here the film “Hitchhicker’s guide to the galaxy” will be recommended to the user by the systems.But this systems generates blunders, for the new genre that the user haven’t watched. Suppose a movie with genre drama is there , then we cannot calculate the approx rating based on users likes/dislikes.

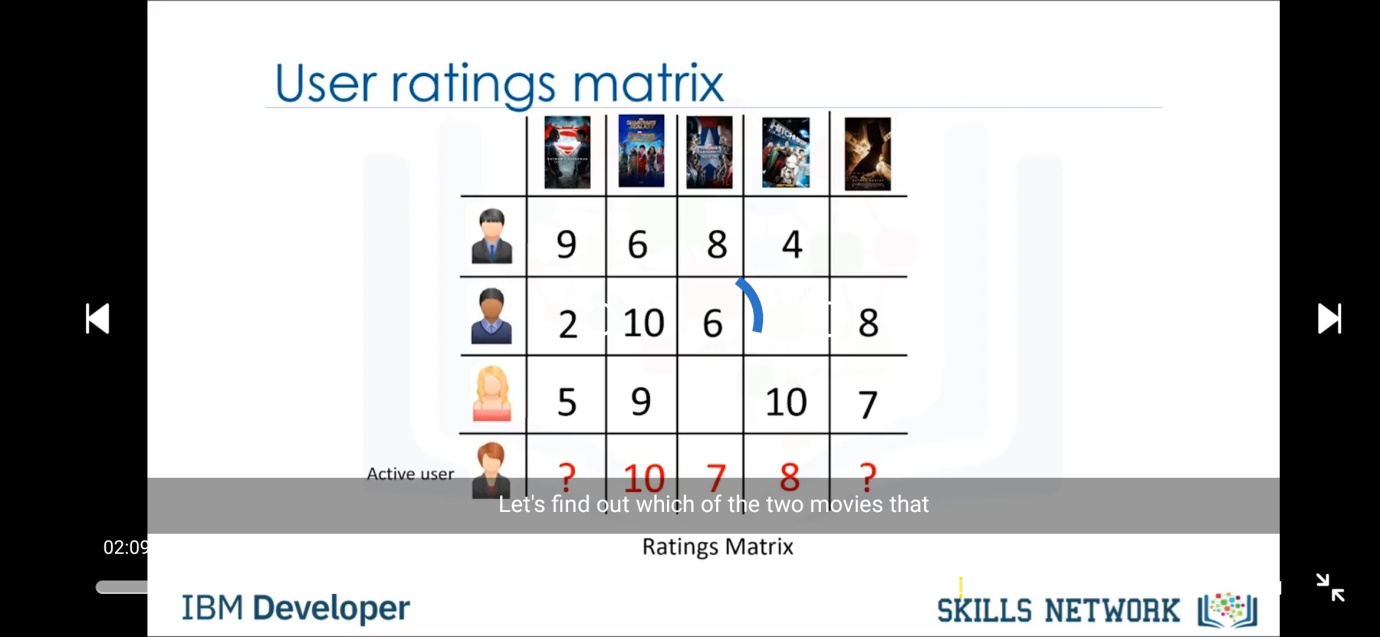
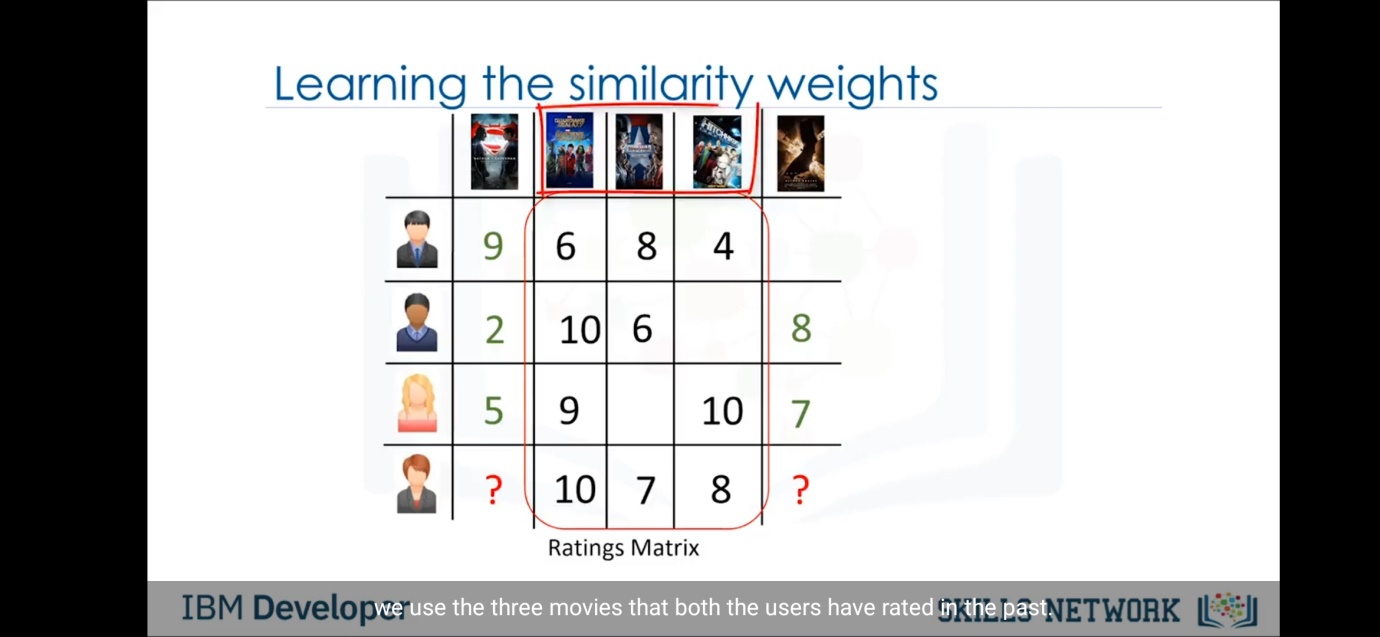
Second type of recommender system is collaborative filtering . In this type the recommendations are based on the other users similar with each others profile , rather than recommending besed on similiar items.

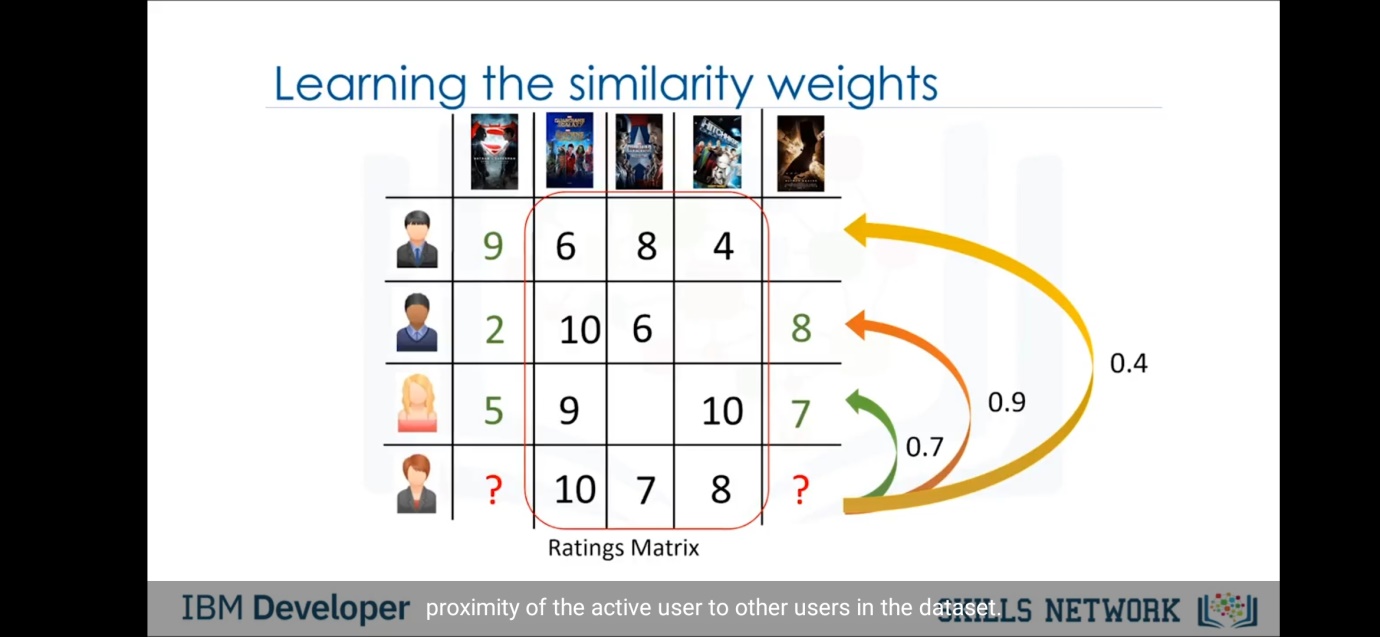
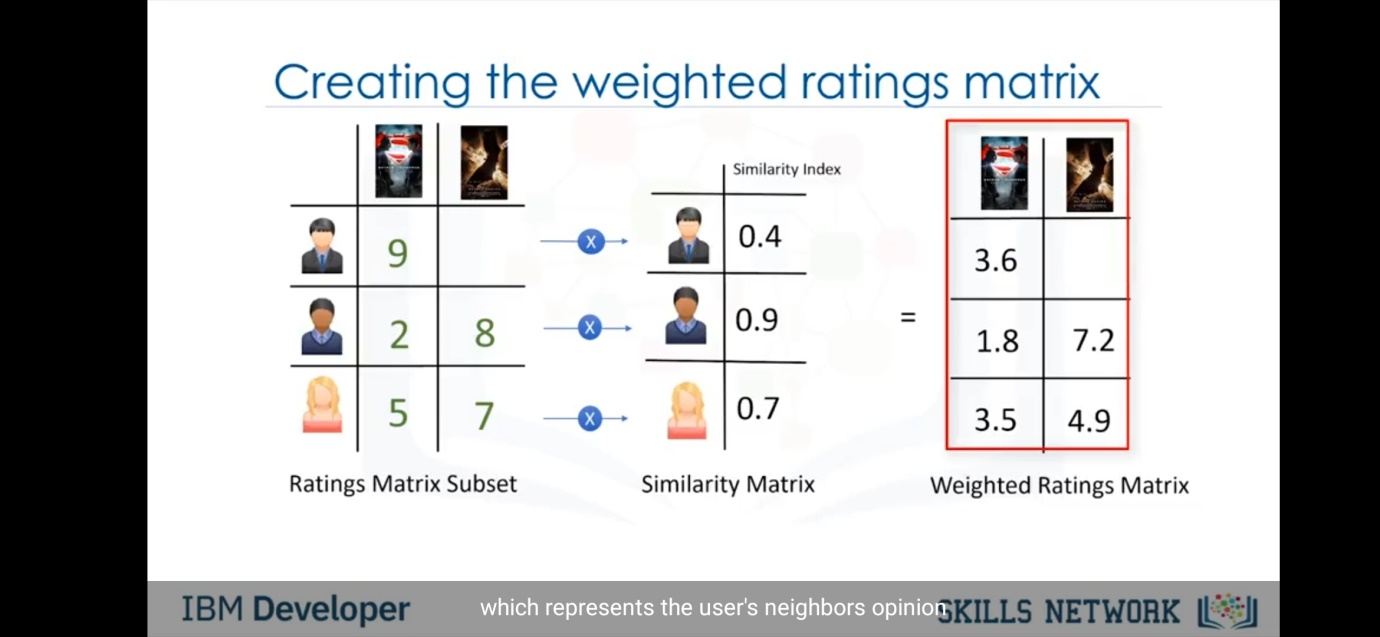
Collaborative filtering are of two types.User based collaborative filtering and item based collaborative filtering.

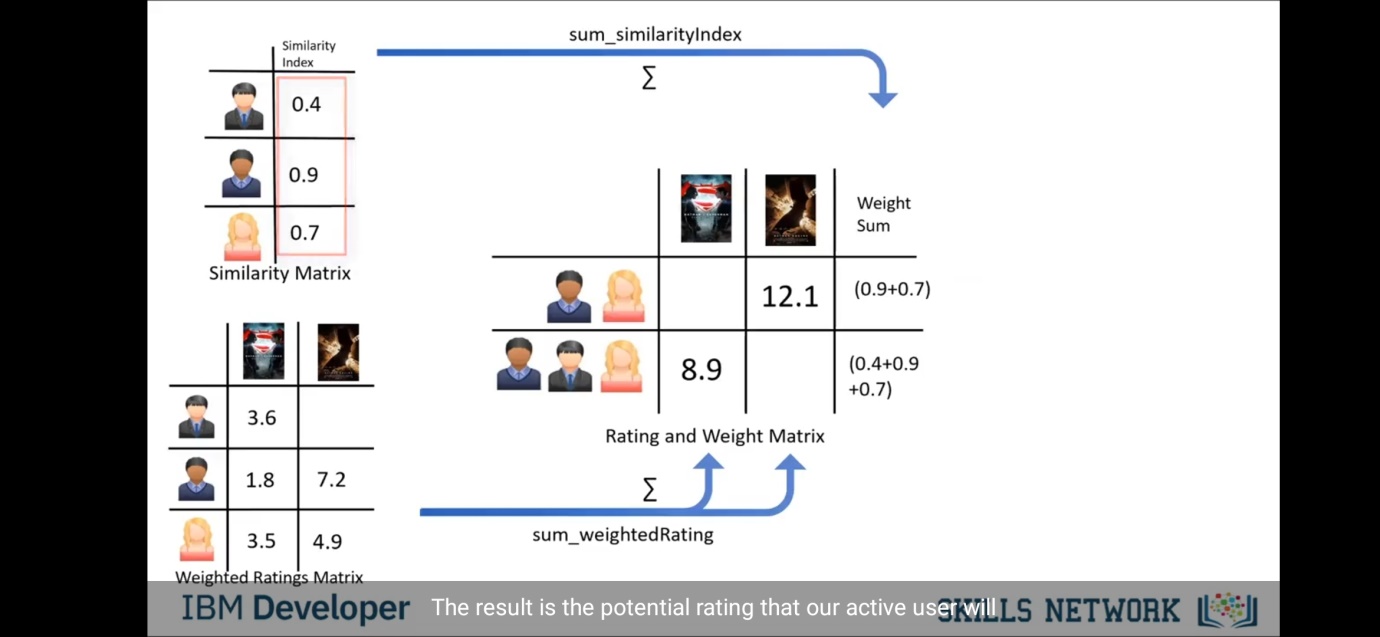
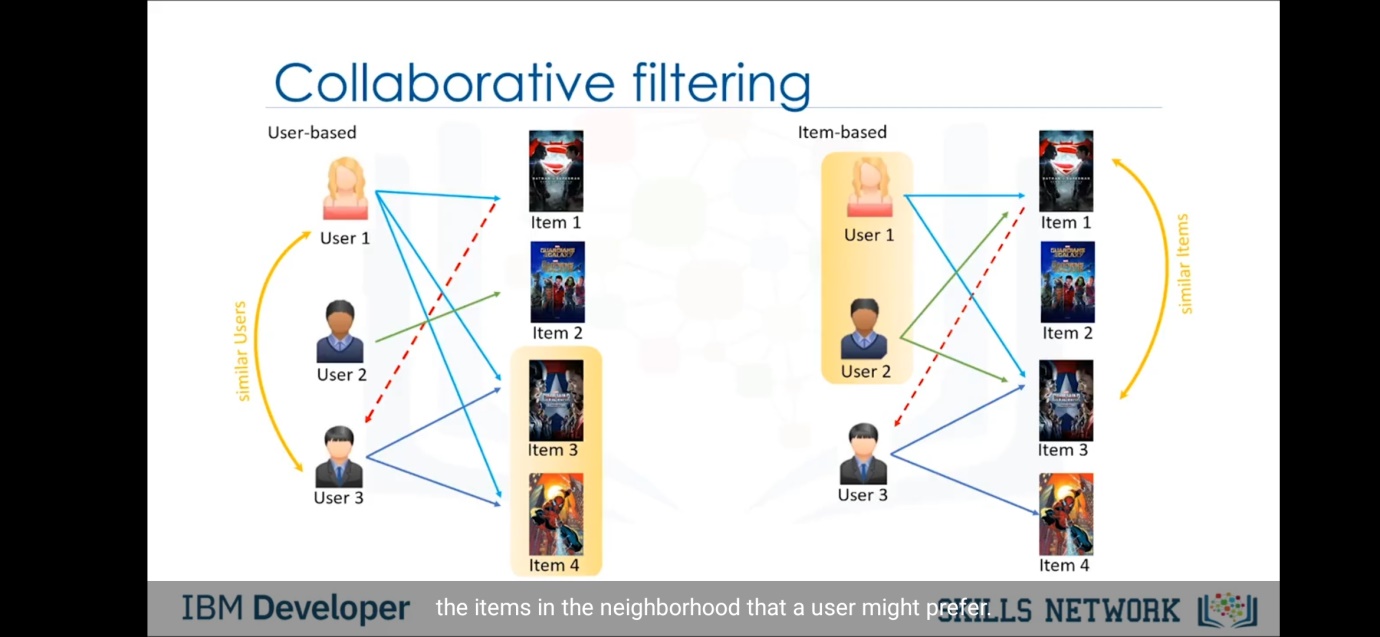
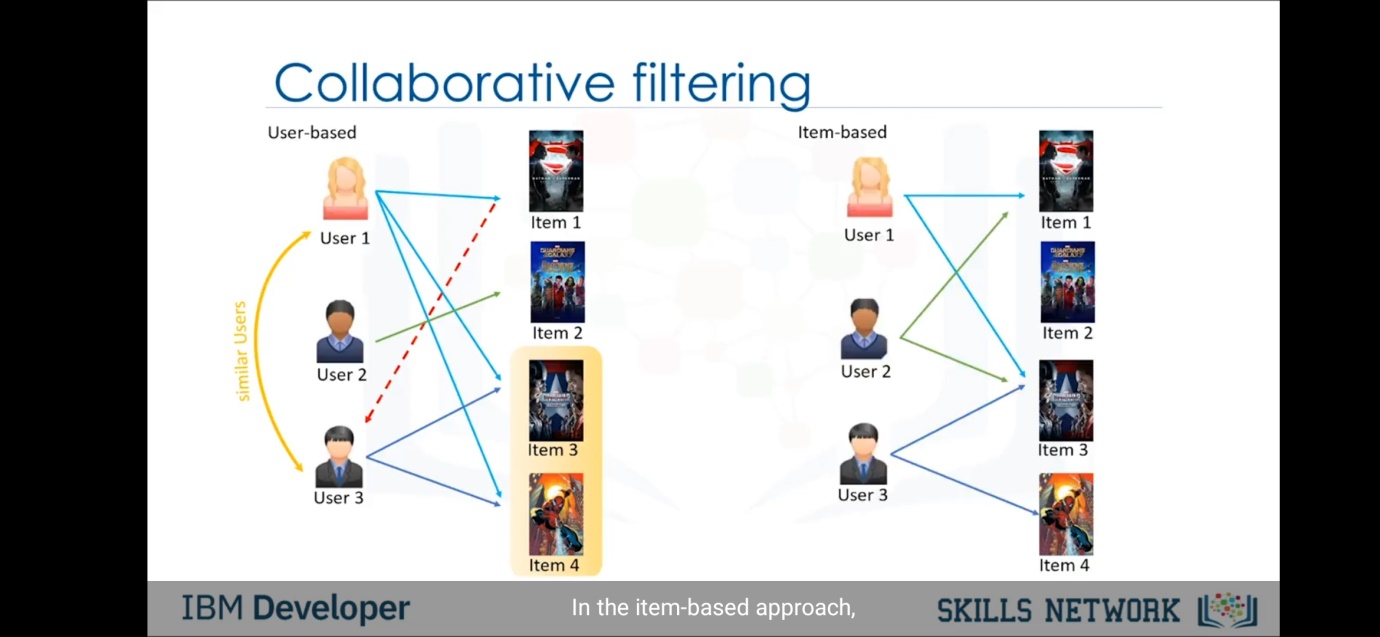
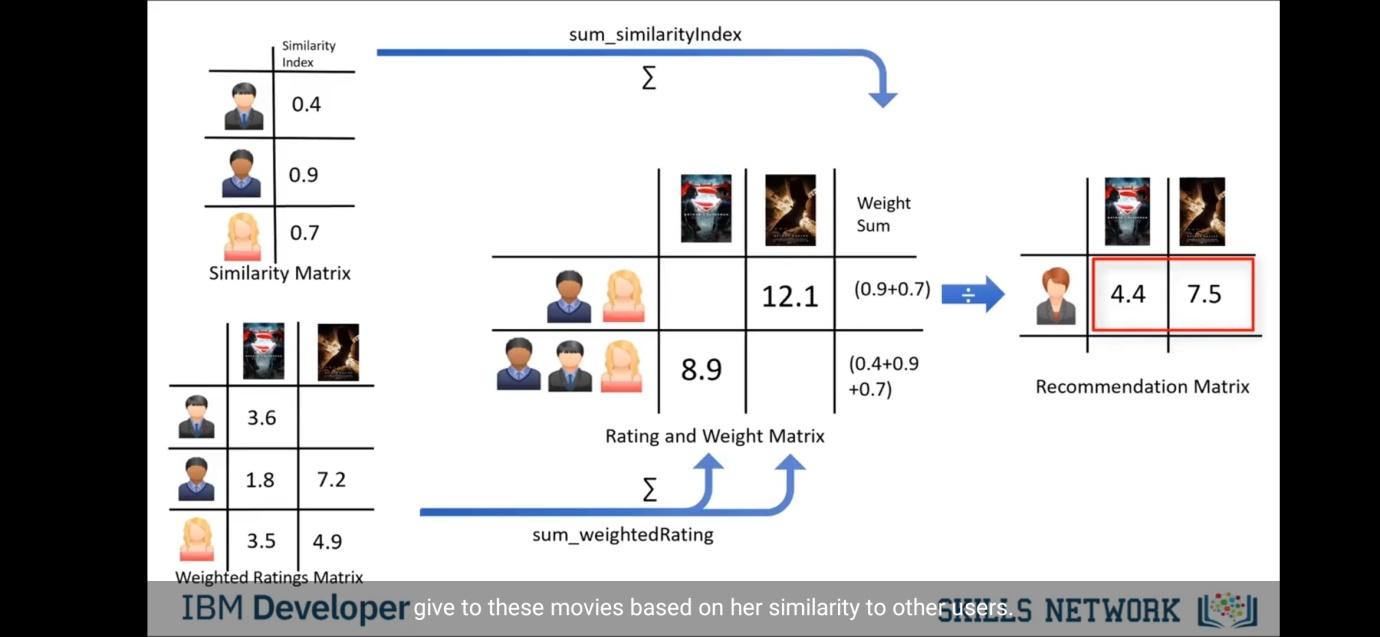
User based collaborative filtering recommends based on users neighbourhood .

Item based filtering recommends based on item similiarity.

First we will study user based collaborative filtering. As shown in the figure if we suppose that there are 2 users , a and b. User a likes 2 movies and user b likes 3 movies out of which 2 movies are common between both users. So the 3rd movie will be recommended by the system to the user a as user a and b likes same 2 movies, so it thinks as user a might actually like the 3rd movie as its like by user b. 

Now we will go through the algorithm of this system.Suppose as given in the figure there are 4 users and some movies atleast one user had watched. Now we mark the 4th user as active user, so we have to find an movie recommendation for that active user based on other users likings.So here the ratings matrix is given in the image as shown.Also we can observe that the active user has not seen the 1st and 5th movie. So to find similiarity between users with respect to the active user, we shall consider only movies wathced by the active user.

Now we can find the similiarity index based on any of the methods like Euclidian distance methods etc.So here in the figure it shows values 0.7 ,0.9 and 0.4 as similarity index of users with respect to the active user.After calculating the similiarity index we will keep it in a matrix called similiarity matrix. Now we will extract the values of movies that are not watched by the active user. So we form a rating matrix of users other than active user with movies not watched by active user. This matrix is called the ratings matrix subset. Now we shall multiply ratings matrix subset with similiarity matrix to obtain weighted ratings matrix.

But here we can see that first user has not seen the movie “Batman v Superman”. So his rating is not present. So now we add all the values of the columns of the weighted ratings matrix and divide it with the user similar matrix value. But here the first film is not watched by the first user so his similiarity is not counted in that movie. So weight sum for 1st movie is (0.9+0.7) because the movie is not watched by user 1. But second movie is watched by all users so weight sum is (0.4+0.9+0.7). After getting rating and weight sum we divide the ratings with the weight sum and get the final recommendation matrix for active user. So here we got values 4.4 and 7.5 for 1st and 2nd movie respectively . So one with more expected rating(here 2nd movie) will be recommended by system to the active user. **Challenges of collaorative filtering:-**

Data Sparsity(Users in general rate only a limited number of items)

Cold start(Difficulty in recommendations to new users or new items)

Scalability(increase in number of users or items)