Recommendation Systems - Part 2

Topics covered so far

- Recommendation Systems
  - Clustering based recommendation system
  - Matrix Estimation with content based
  - Matrix Estimation over time
  - Hybrid recommendation systems

Discussion questions

1. What are content based recommendation systems?
2. How do you build clustering based recommendation systems?
3. What are hybrid recommendation systems? Why are they useful?
4. How do you forecast the time series data using matrix estimation techniques?

Content based recommendation systems

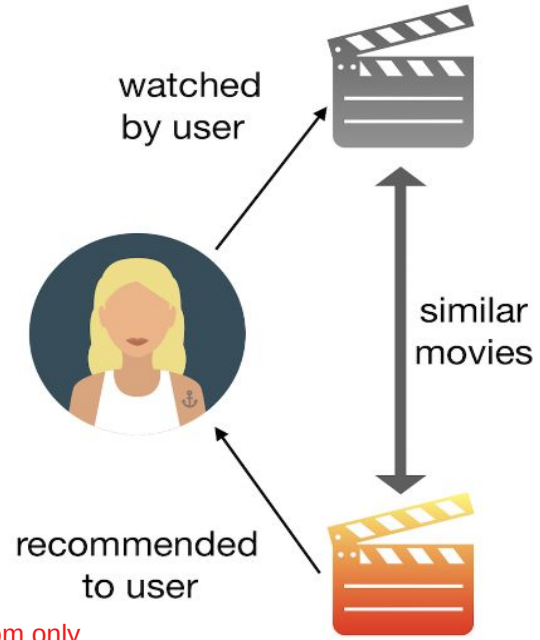
- Content based recommenders work solely with the past interactions of a given user and do not take other users into consideration
- The main idea behind content based recommendation system is that **if a user likes an item, then he/she will also like a "similar" item**
- Uses information of all products and preferences of just that one user

Advantages

- Content-based recommender systems don't require a lot of user data
- Does not suffer from cold start
- Less expensive to build and maintain

Challenges

- Unavailability of features which explain Items and user preferences
- Recommendations will likely be direct substitutes, instead of complements, of the item the user interacted with
- Less dynamic



Source: towardsdatascience

Clustering based recommendation systems

We can also apply unsupervised methods (clustering) to build recommendation systems. Each cluster would be assigned to typical preferences based on preferences of customers who belong to the cluster. Customers within each cluster would receive recommendations computed at the cluster level.

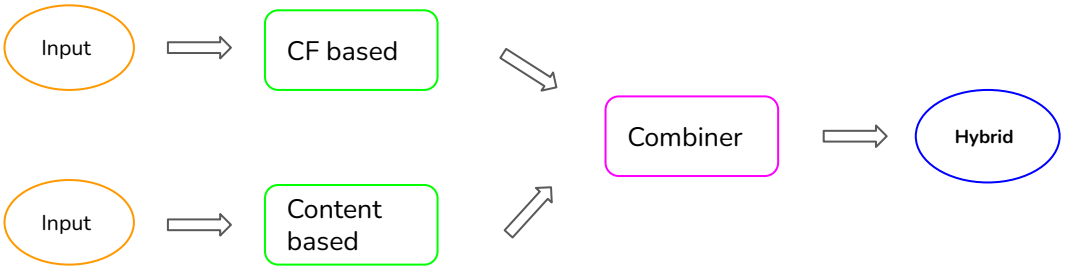
It works in the followings steps:

1. Compute similarity between each pair of users
2. Obtain representation of each user in low dimensional space
3. Perform the K-means clustering algorithm and find the number of clusters, i.e., K

This image shows the clusters of users who have similar preferences so that they will receive the same product recommendations.

Hybrid recommendation systems

- Hybrid recommender system is a special type of recommender system that combines different methods, generally, content and collaborative filtering methods
- Hybrid approaches can be implemented in several ways: by making content-based and collaborative-based predictions separately and then combining them; by adding content-based capabilities to a collaborative-based approach (and vice-versa)
- These methods can also be used to overcome some of the following problems in recommender systems:
  - Avoid the problem of recommending only similar products
  - Overcome the problem of false nearest neighbors
  - More personalized recommendations and specific to the user needs
  - Computationally faster system than the traditional recommendation system algorithms



Matrix estimation on time series data

- Matrix estimation can be used to fill missing values and forecast the time series data.
- Let's say you are given the following time series: X(1), X(2), ?, ?, X(5).....X(L)..... X(T). This matrix can be represented in the form of matrix as given below.
- Take some value L, say 3, and take first L values and put them as a column in a matrix. Keep repeating this for T values. After that, we can apply different matrix estimation methods to fill in the missing values, and then forecast next L values in the dataset **(X(T+1), X(T+2), X(T+3))**.

X(1)	?	.	.	.	.	.	.	X(T+1)
X(2)	X(5)	.	.	.	.	.	.	X(T+2)
?		.	.	.	.	.	X(T)	X(T+3)

- The good things about this process is, unlike the traditional forecasting methods which assumes stationarity, we don't have any assumptions here, and can apply this method for forecasting.

Case Study

Great Learning

Happy Learning !