Q1. Agglomerative Hierarchical Clustering.

<u>Hierarchical Methods:</u>

This method creates a hierarchical decomposition of the given set of data objects. We can classify hierarchical methods on the basis of how the hierarchical decomposition is formed. There are two approaches here –

- **Agglomerative Approach**
- **Divisive Approach**

In hierarchical clustering, a treelike cluster structure (dendrogram) is created through recursive partitioning (divisive methods) or combining (agglomerative) of existing clusters.

• Agglomerative Approach This approach is also known as the bottom-up approach. In this, we start with each object forming a separate cluster. It keeps on merging the objects or clusters that are close to one another. It keep on doing so until all of the clusters are merged into one or until the termination condition holds.

<u>Steps for Hierarchical Clustering – Agglomerative approach:</u>

1. Compute distance matrix from object features.
2. Set each object as a independent cluster.(if there are 5 objects , then there will be 5 clusters)
3. Iterate until number of cluster is equal to 1 A. Merge two closest clusters B. Update distance matrix

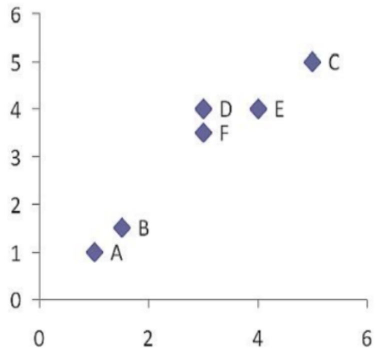
Assume we have six objects A,B,C,D,E,F each having two attribute X1 and X2

Distance between two objects is calculated using Euclidian distance formula using Their attributes X1 and X2.

For example distance between A and B can be calculated as:

$$d(A,B) = \sqrt{(X_{A1} - X_{B1})^2 + (X_{A2} - X_{B2})^2}$$

Object	X1	X2
A	1	1
B	1.5	1.5
C	5	5
D	3	4
E	4	4
F	3	3.5



For example, distance between object A = (1, 1) and B = (1.5, 1.5) is computed as

$$d_{AB} = \left((1-1.5)^2 + (1-1.5)^2\right)^{\frac{1}{2}} = \sqrt{\tfrac{1}{2}} = 0.7071$$

Another example of distance between object D = (3, 4) and F = (3, 3.5) is calculated as

$$d_{DF} = \left((3-3)^2 + (4-3.5)^2\right)^{\frac{1}{2}} = 0.5$$

Q2. HOW SIMILARITY IS MEASURED IN CLUSTERING TECHNIQUE

Clustering is a technique to group objects based on distance or similarity

Clustering algorithms seek to segment the entire data set into relatively homogeneous subgroups or clusters, where

- The similarity of the records within the cluster is maximized, and
- The similarity to records outside this cluster is minimized.

How to measure similarity:

For measuring similarity Distance metric is used.

Most common distance metric is Euclidean Distance. Other Distances can also be used.

Distance functions

Euclidean $\sqrt{\sum_{i=1}^{k}(x_i - y_i)^2}$

Manhattan $\sum_{i=1}^{k}|x_i - y_i|$

Minkowski $\left(\sum_{i=1}^{k}(|x_i - y_i|)^q\right)^{1/q}$

where x = x1, x2, … , xm, and y = y1 , y2, … , ym represent the m attribute values of two records

Q3. How centroid is calculated for multiple att. Data in clustering.

K- mean algorithm

Step 1: Select number of clusters k the data set should be partitioned into.

Step 2: Randomly assign k records to be the initial cluster(Usually first k record are assigned to K clusers)

Step3: Calculate centroid of the cluster.

Step 4: For each record, find the nearest cluster center and add the record to that cluster.

Step5: For each of the k clusters, find the cluster centroid, and update the location of each cluster center to the new value of the centroid.

Step 6: Repeat steps 4–5 until convergence or termination(centroid do not change).

Centroid of the cluster is the mean value of the elements in that cluster

(For Example : Clustering Pdf Page No: 10)

Q4. Collaborative Filter based Recommender System

Neighbourhood-based recommendation engines:

Collaborative Filtering

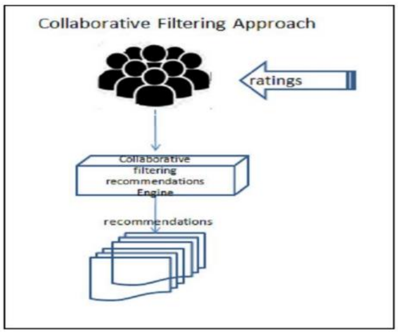
Neighbourhood-based recommender systems considers the preferences or likes of the user community or users of the neighbourhood of an active user before making suggestions or recommendations to the active user.

The idea for neighbourhood-based recommenders is very simple: given the ratings of a user, find all the users similar to the active user who had similar preferences in the past and then make predictions regarding all unknown products that the active user has not rated but are being rated in by his neighbourhood

Types:

- **User-based collaborative filtering.**
- **Item-based collaborative filtering.**

Collaborative filtering is the process of filtering for information or patterns using techniques involving collaboration among multiple agents, viewpoints, data sources, etc



User-based collaborative filtering

User based collaborative filtering first finds out the similarity between the active user (the user needing the recommendation) and other users.

Identifies the similar users based on Euclidian distance or correlation coefficient.

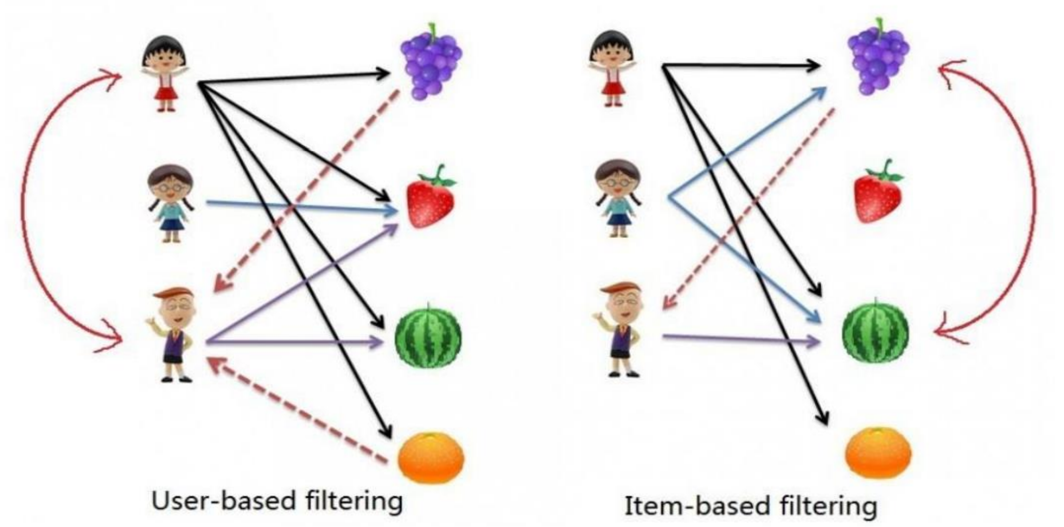
Recommend the products that has not rated/purchased by active user but rated/purchased by similar/ nearest users.

Example: Movie Recommendation

Item-based collaborative filtering

Item-based collaborative filtering recommender systems, unlike user-based collaborative filtering, we use similarity between items instead of similarity between users.

The basic intuition for itembased recommender systems is that if a user liked item A in the past they might like item B, which is similar to item A:



Q5. EQUATION FOR PARAMETER ESTIMATION IN MULTILAYER NEURAL NETWORK USING BACK PROPAGATION TECHNIQUE.

https://towardsdatascience.com/understanding-backpropagation-algorithm-7bb3aa2f95fd

Q6. RESIGN ARTIFICIAL NEURON FOR EX-OR OPERATION & OR OPERATION.

https://towardsdatascience.com/implementing-the-xor-gate-using-backpropagation-in-neural-networks-c1f255b4f20d

Q7. TECHNIQUE FOR SERIES PRODUCTION

Q8. K-MEAN CLUSTERING ALGORITHOM.

K-mean Clustering Algorithm

K-Means clustering intends to partition n objects into k clusters in which each object belongs to the cluster with the nearest mean.

This method produces exactly k different clusters of greatest possible distinction.

The best number of clusters k leading to the greatest separation (distance) is not known as a priori and must be computed from the data.

Follow Que Ans for Steps .....