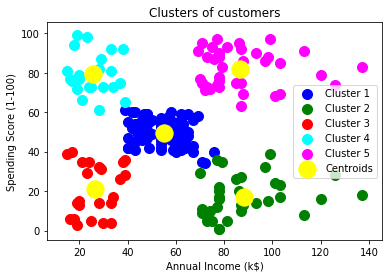
**UNSUPERVISED LEARNING**

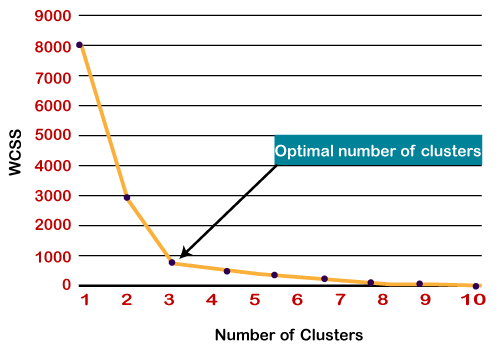
 INSTEAD OF GIVING THIS INPUT FEATURES GIVES THESE OUTPUT TO THE ALGORITHM. WE ARE GIVING ONLY THE INPUT FEATURES THEN THE ALGORITHM FINDS THE OUTPUT (CLUSTERS AND ASSOCIATIONS).

TWO TYPES

1.CLUSTERING (GROUPING THE DATA POINTS INTO CLUSTERS)

2.ASSOCIATION (FINDING RELATIONSHIP BETWEEN DATAPOINTS)

1.FAMOUS CLUSTERING MODELS



1**. K-means clustering: (**partitions the data into a pre-defined number of clusters)

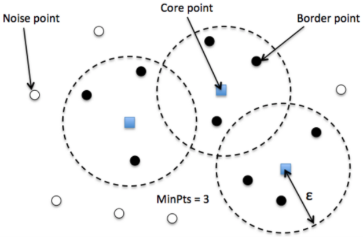
If you give **k=2** ,it randomly selects two data points called **centeroid** and measure the

distance between datapoints and calculate the (MSE) error and change the centeroids

iteratively .The only manual thing is find the k using **elbow hand** method.(the image

shows k==3 is the best fit)(min max scalar optimization important)

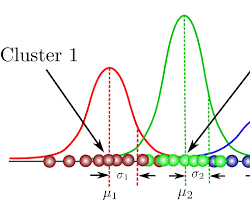
2**. DBSCAN (Density-Based Spatial Clustering of Applications with Noise):**

Does not require the number of clusters to be known beforehand. It works by finding

clusters of high-density data points that are separated by low-density regions.

Pros: DBSCAN is a good choice if the **data contains noise and outliers**.

Optimization techniques:1. flat/uniform 2. Gaussian



**3. Gaussian mixture models (GMM):**

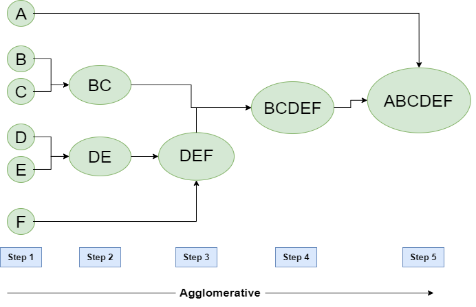
**Clustering and finding the output in the probabilistic method.**

**Each clusters have one mean & variance. The side pictures show**

**How its works.** GMM is a good choice if the data is probabilistic in nature.

**(Clusters are classified by standard deviation and covariance)**

4. **Agglomerative hierarchical clustering:**

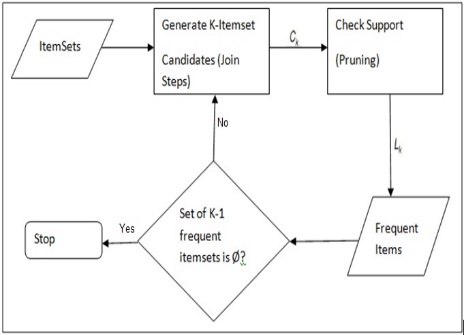
This is a hierarchical clustering algorithm that starts by assigning each

data point to its own cluster. It then iteratively merges the clusters that are most

similar to each other until there is only one cluster left. Agglomerative hierarchical

clustering is a good choice if the data needs to be clustered in a hierarchical manner.

Some sort of procedures followed to build the network

2.FAMOUS ASSOCIATION FINDING MODELS

**1.Apriori Algorithm:**

Find the relation between datapoints how much it merged together and the

Weight of the association (it used by huge markets data).SKlearn could not have

Any Specific library but it used rarely. optimization technique plays main roll (**hash**

**Table and bit vector**)