**Course: Advanced Bioinformatics**

**Module title: Clustering**

**Module no. : 184**

Clustering is a process of organizing objects into groups whose members are similar in some way.

A cluster is therefore a collection of objects which are similar between them and are dissimilar to the objects belonging to other clusters

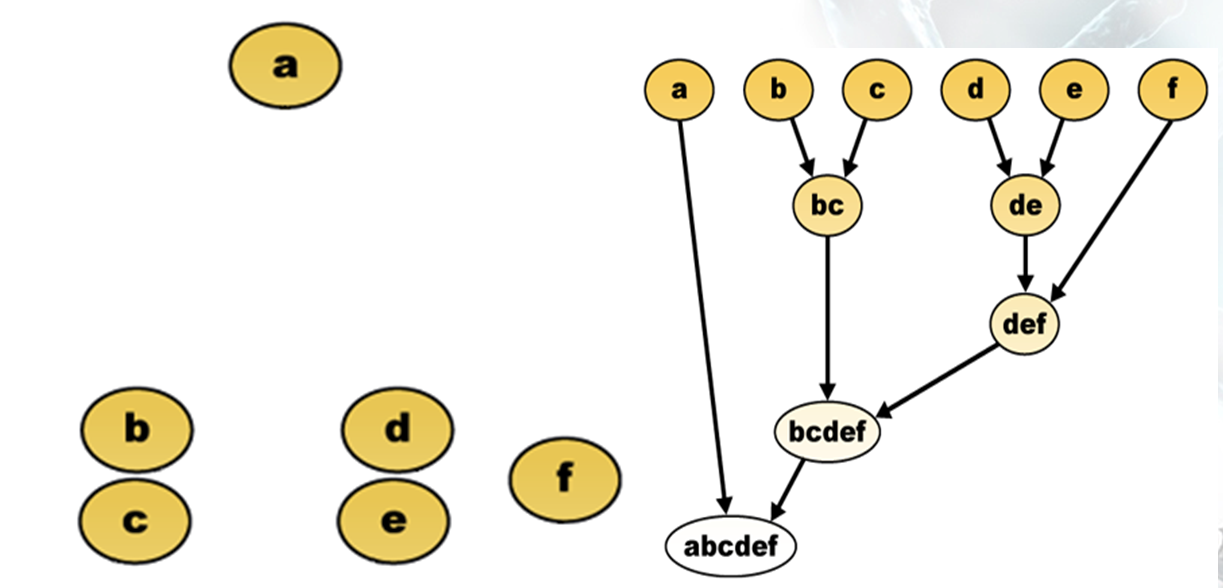
Useful in data concept construction

Unsupervised learning process

Hierarchical

Agglomerative general algorithm

* Find the 2 closest objects and merge them into a cluster
* Find and merge the next two closest points, where a point is either an individual object or a cluster of objects
* If more than one cluster remains, return to step 2



Applications:

* For administrative purposes
* Hospital activity and performance
* Used for researchers
* Data used by physician
* For laboratory use etc.