

Hierarchical Clustering

Consider the vertebrate dataset as below (ignore class IDs)

	Name	Warm-blooded	Gives Birth	Aquatic Creature	Aerial Creature	Has Legs	Hibernates	Class
0	human	1	1	0	0	1	0	mammals
1	python	0	0	0	0	0	1	reptiles
2	salmon	0	0	1	0	0	0	fishes
3	whale	1	1	1	0	0	0	mammals
4	frog	0	0	1	0	1	1	amphibians
5	komodo	0	0	0	0	1	0	reptiles
6	bat	1	1	0	1	1	1	mammals
7	pigeon	1	0	0	1	1	0	birds
8	cat	1	1	0	0	1	0	mammals
9	leopard shark	0	1	1	0	0	0	fishes
10	turtle	0	0	1	0	1	0	reptiles
11	penguin	1	0	1	0	1	0	birds
12	porcupine	1	1	0	0	1	1	mammals
13	eel	0	0	1	0	0	0	fishes
14	salamander	0	0	1	0	1	1	amphibians

Use (1) single link (MIN), (2) complete link (MAX), and (3) group average for clustering. Plot dendrogram for each method.

Density Based Clustering

Consider Mydata.mat file, which includes the data points in 2-D space.

- 1- For $k=10$, plot the k -dist of the points and determine a good value for Eps. (Hint: For all points, find the distance of point to its k^{th} nearest neighbor, sort the distance values and plot)
- 2- Determine the core points, the border points and the noise points
- 3- Use DBSCAN algorithm for clustering
- 4- Plot the data