

Lab on data clustering

Thomas Bonald

MS Big Data
2018 – 2019

3 topics

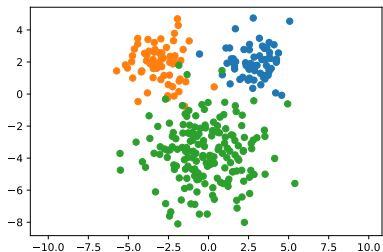
Do 1, 2 or 3 among:

1. *k*-means vs. **Gaussian mixture**
2. **Image segmentation** by *k*-means
3. **Hierarchical clustering** and best cuts of dendrograms

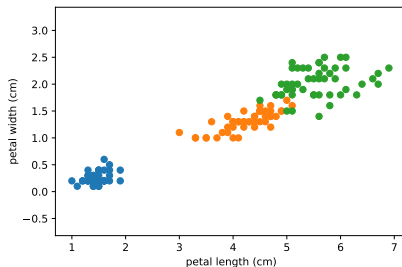
See the Python notebooks

1. k -means vs. Gaussian mixture

Gaussian mixture model



Iris dataset



2. Image segmentation by k -means

Data



k -means ($k = 4$)



Mask



Segmentation



3. Hierarchical clustering

Karate club

