Pattern Analysis Thursday Meeting
Juliane Hoffmann

Worksheet 3: Thursday, May 22, 2025

Task 1: Mean Shift Clustering

(a) What is the idea behind the mean shift clustering?

(b) What are the kernel constraints?

(c) What are the steps of the mean shift algorithm?

Task 2: Kernel Selection for Mean Shift Clustering

(a) What role does the bandwidth of the kernel play in Mean Shift Clustering?

(b) How does increasing the bandwidth affect the clustering result?

(c) What do the clustering results look like when using a smaller bandwidth?

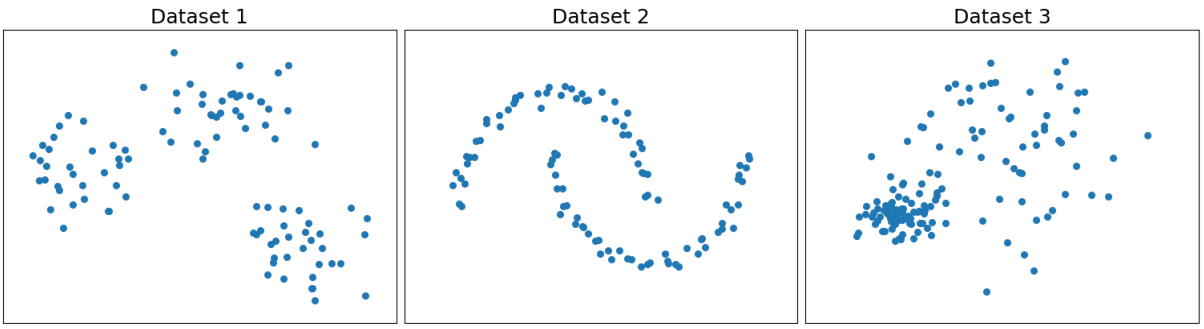
(d) What happens to the clustering process if the kernel is not radially symmetric (going against the defined constraints of the kernel function)?

Task 3: K-means Model Selection

What does the optimality criterion given in Equation 1 aim to achieve?

$$k^* = \underset{k}{\operatorname{argmin}}\{k \mid G(k) \geq G(k+1) - s'_{k+1}\} \tag{1}$$

Task 4: Algorithm Selection



For each of the three datasets shown above:

(a) Describe the shape and distribution of the data

(b) Decide which clustering algorithm among the following are most suitable and explain why:

- K-means
- Gaussian Mixture Model (GMM)
- Mean Shift

(c) Justify why the other methods might fail