

**Deadline: Th. Jan. 24, 10:00 am** Drop your printed or legible handwritten submissions into the boxes at Samelsonplatz, or upload them as `.pdf` or `.ipynb` files onto the LearnWeb.

**Exercise 1** (K-Means - 10 points).

Given the data:

$x_1$	$x_2$
2	2
2	1
0.9	0.9
0	0
-1	0
-1.2	0
-1	-1

Perform two iterations of the K-means clustering algorithm for  $K = 2$  and  $K = 3$ . Use  $\mu_1 = 0$  in each case.

**Exercise 2** (Gaussian Mixture Model - 2+2 points).

1. (2) Explain why K-means clustering with the euclidean distance metric is just a special case of a GMM.
2. (2\*) Explain why K-means clustering with the Manhattan distance metric is not a special case of a GMM.

**Exercise 3** (Hierarchical Models - 8 points). Consider the data set

$$X = \{(k, l) \in \mathbb{R}^2 \mid k \in \{-3.5, -2.5, \dots, +2.5, +3.5\}, l \in \{-1, +1\}\}$$

1. (3) Draw the data-set. What are the partitions resulting from K-means algorithm with  $K = 2$  at convergence?
2. (5) Apply agglomerative hierarchical clustering using the single link distance. Compare the level 1 partitions with the result of the  $K$ -means algorithm

Note: If  $\min_n f(n)$  is found for multiple  $n$ , choose one at random.