WRITE FIRST NAME, LAST NAME, AND ID NUMBER ("MATRICOLA") ON YOUR ASSIGNMENT. TIME: 1.5 hours.

FIRST NAME: .....

LAST NAME: .....

ID NUMBER: .....

Exercise 1 [6 points]

1. Introduce the main assumptions typically made in a learning problem.
2. Define the *regression* problem and introduce:
   - (at least) two model classes that can be used in a regression framework
   - a criterion that can be used to chose the "best" model once the model class is fixed
   - a criterion to chose the "best" model class among a few alternatives.

[Solution: Exercise 1]

[Solution: Exercise 1]

[Solution: Exercise 1]

Exercise 2 [6 points]

1. Describe the principle of Support Vector Machines for (binary) classification, and formulate the optimization problems for Hard and Soft SVM

2. How can SVM be extended for non-linear classification? Define the concept of Kernel and highlight how the decision boundary depends on the Kernel function and the training data

---

[Solution: Exercise 2]

[Solution: Exercise 2]

[Solution: Exercise 2]

Exercise 3 [6 points]

Consider a linear regression problem with predictor function

$$h_w(x) = x^\top w \quad x \in \mathbb{R}^2 \quad w \in \mathbb{R}^2$$

and measurement model

$$y_i = h_w(x_i) + e_i \qquad e_i \sim \mathcal{N}(0, \sigma^2), \quad i \in [1, .., m]$$

We assume that some prior knowledge on w is available, which can be encoded postulating that w is a Gaussian random vector with mean w_o and variance γI, $\gamma > 0$. We also assume that the measurement noises e_i are independent of w.

1. Compute the distribution of the measurement vector $Y := [y_1, y_2, .., y_m]^\top \in \mathbb{R}^m$. Assuming γ is known, discuss how the distribution of Y can be used to compute an estimator of the noise variance σ^2.

2. Compute a closed form expression for the MAP (Maximum a Posterior) estimator $\hat{w}_{MAP}$ of w.

[Solution: Exercise 3]

[Solution: Exercise 3]

[Solution: Exercise 3]

Exercise 4 [6 points]

1. Describe the K-means algorithm for clustering.
2. Assuming you are given a dataset $\{x_i\}_{i=1,..,M}$ and you DON'T know the "right" number of clusters. Describe a possible procedure that allows to estimate both the cluster centers as well as the number of clusters.

[Solution: Exercise 4]

[Solution: Exercise 4]

[Solution: Exercise 4]