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. Define the regression problem and describe *at least* 3 model classes we have encountered and used to solve regression problems
2. Discuss the bias-variance tradeoff and how it can be handled.

[Solution: Exercise 1]

[Solution: Exercise 1]

Exercise 2 [6 points]

Assume you are given samples $\{x_i\}_{i=1,..,m}$, $x_i \in \mathbb{R}$ i.i.d. from and unknown distribution $\mathcal{D}$. Your objective is to estimate the probability density function $p_{\mathbf{x}}(x)$. To do so, you make the assumption that $p_{\mathbf{x}}(x)$ can be written in the folwing form:

$$p_{\mathbf{x}}(x) = \sum_{i=1}^{n} \alpha_i p_i(x)$$

where n is known and fixed, α_i are real (and non-negative) parameters (unknown) and

$$p_i(x) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{1}{2\sigma_i^2}(x-\mu_i)^2}$$

are parametrixed by μ_i (unknown) and $\sigma_i > 0$ (unknown).

1. Describe an algorithm to estimate the unknown parameters
2. Is this problem related to any of the problems studied in class? If so, which and how?

[Solution: Exercise 2]

[Solution: Exercise 2]

Exercise 3 [6 points]

Kernels are used in several machine learning algorithms, among which Gaussian regression (regression using Gaussian Processes) and SVMs.

1. Starting from the strucure of the solution, discuss the influence that the kernel has on the solution, and point out how SVM and Gaussian regression differ.

2. In the contex of Gaussian Regression, assume that the kernel function is parametrized by a scale parameter, i.e.

$$K_\lambda(x, x') = \lambda K(x, x') \quad \lambda \geq 0$$

   where $K(x, x')$ is a fixed (positive definite) function. Describe (at least) one procedure to estimate λ from data.

[Solution: Exercise 3]

[Solution: Exercise 3]

Exercise 4 [6 points]

Consider the soft (linear) SVM classification problem and let c be the constraint on the slack variables, i.e.

$$\sum_{i=1}^{m} \xi_i \leq C$$

1. Formulate the optimization problem to find the separating hyperplane
2. Discuss the relation between C and the margin.

[Solution: Exercise 4]

[Solution: Exercise 4]