

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

FIRST NAME:

LAST NAME:

ID NUMBER:

Question 1 [6 points]

1. Introduce the classification problem in machine learning.
2. With reference to linear binary classification:
 - briefly describe the perceptron algorithm and the logistic regression model
 - point out the main difference among the two.

[Solution: Question 1]

[Solution: Question 1]

Question 2 [6 points]

Assume you are given data $\{x_i\}_{i=1,\dots,m}$, $x_i \in \mathbb{R}^d$, that you would like to represent (approximately) using a small number $k < d$ of “parameters”.

1. Describe a linear technique to perform this reduction and, for a given number of reduced parameter,s find a closed form expression for this approximation.
2. Discuss the relation between the number of parameters and the (average) approximation error.

[Solution: Question 2]

[Solution: Question 2]

Question 3 [6 points]

Give the definition of on-average-replace-one-stable algorithms and discuss the relation between (OAROS) stability and overfitting.

[Solution: Question 3]

[Solution: Question 3]

Question 4 [6 points]

Assume you have data (x_i, y_i) , $i = 1, \dots, m$, $x_i \in \mathbb{R}^d$, $y_i \in \mathbb{R}$ and you would like to learn a (non-linear) model that approximates the outputs y_i with their predictions

$$\hat{y} := h(x)$$

You would like to do so under the assumption that the labels y_i have been generated by the (unknown) model

$$y_i = f(x_i) + e_i$$

where e_i are i.i.d. zero mean Gaussian noises with unknown variance σ^2 .

1. State this learning problem in the context of Gaussian Processes (i.e. assume $f(x)$ is a realization from a Gaussian Process with zero mean and known covariance function) and write the corresponding estimate $h(x)$.
2. In this framework, describe a procedure to estimate the unknown noise variance σ^2 .

[Solution: Question 4]

[Solution: Question 4]