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 main ingredients of a (supervised) learning problem, together with the typical assumptions made.

2. Formulate the PAC learning problem under the realizability assumption and provide guarantees on the sample complexity under the finite model class assumption.

[Solution: Question 1]

[Solution: Question 1]

Question 2 [6 points]

With reference to a binary classification problem:

1. Starting from the definition of linear separation boubndary, derive the SVM formulation (hard case) and describe how it can be extended for non-linearly seperable data

2. Explain how kernel SVMs can be formulated by introducing a feature map $z := \Phi(x)$.

[Solution: Question 2]

[Solution: Question 2]

Question 3 [6 points]

Formulate the clustering problem, introduce a suitable cost function J to be optimized and describe the k-means iterative algorithm. In particular:

1. Prove that both the updates for labels and centroids decrease the value of the cost J
2. Is it possible to prove that this algorithm converges to the global optimum of J?

[Solution: Question 3]

[Solution: Question 3]

Question 4 [6 points]

Inthe context of regularization as tool to deal with overfitting.

1. Define the cost J for an ℓ^2- regularized linear regression problem (ridge-regresion) with regularization parameter λ. Denoting with $\hat{h}_R(\lambda)$ the optimal model as a function of the regularization parameter λ and with $\hat{h}_S$ the ERM, draw a typical behaviour of $L_{\mathcal{D}}(\hat{h}_R(\lambda))$, $L_{\mathcal{S}}(\hat{h}_R(\lambda))$ and $L_{\mathcal{S}}(\hat{h}_S)$

2. Discuss *two* alternative methods to tune the regularization parameter λ.

[Solution: Question 4]

[Solution: Question 4]