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

FIRST NAME: .....

LAST NAME: .....

ID NUMBER: .....

Exercise 1 [8 points]

1. Describe the general framework of regression.

2. Write the optimization problem for linear regression with squared loss, and derive a solution for the situation in which $\mathbf{X}^T\mathbf{X}$ is not invertible.

Exercise 2 [8 points]

Consider a binary classification problem with 0-1 loss.

1. Given an hypothesis set $\mathcal{H}$, provide the definition of empirical error and true risk (generalization error) for an arbitrary hypothesis $h \in \mathcal{H}$. Assuming the hypothesis set $\mathcal{H}$ contains only a finite number of elements, provide an expression to bound the generalization error in terms of the emprical error and the sample size.

2. Assume that you have n *finite* hypothesis sets, denoted by $\mathcal{H}_i, i = 1, 2, \ldots, n$ (with finite cardinality $|\mathcal{H}_i|$). Describe one strategy to choose a good hypothesis set $\mathcal{H}_i$ and a good model $\hat{h}_i \in \mathcal{H}_i$.

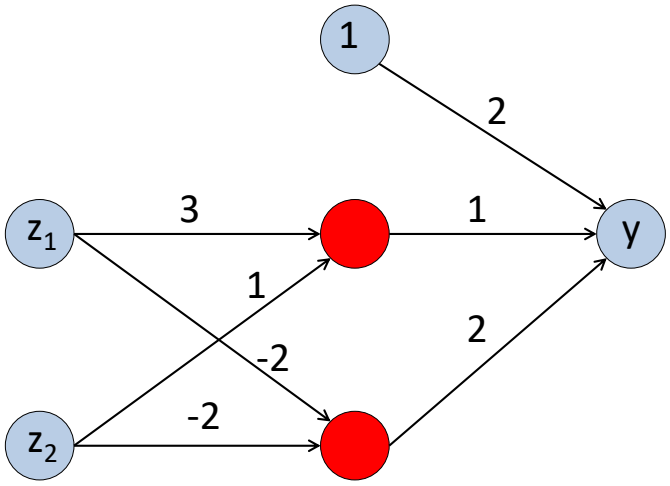
Exercise 3 [8 points]

1. Consider the neural network in the figure below and assume that a Rectified Linear Unit (ReLU) activation function

$$\sigma(x) = \max(0, x)$$

is used for all neurons. Compute the value of the output y when the input $\mathbf{z}$ is the vector $[z_1 \;\; z_2] = [2 \;\; -1]$; (only red nodes include activation functions).

2. Discuss how the neural network can be trained using SGD and the relation with the backpropagation algorithm (only the main structure, details of the derivation are not required).



Exercise 4 [8 points]

Consider the problem of dimensionality reduction.

1. Present a dimensionality reduction technique a briefly explain how it works.

2. Assume you need to classify the data in the Figure below using a machine learning algorithm. Describe what happens when you apply your dimensionality reduction technique to the data in the Figure in order to convert them from the bi-dimensional representation in the Figure to a one dimensional representation (you can use a drawing to show the transformation). Does your dimensionality reduction approach allow to use a simpler classifier in order to recognize the two classes? Provide a brief explanation.