

Master's degree in Artificial Intelligence and Computer Science

OPTIMIZATION FOR MACHINE LEARNING - 6 CFU

Project N. 6

Given the dataset **dataset6.mat**, where X is the matrix whose rows contain the points to be classified and y is the array of the corresponding class labels, perform a **Support Vector Machine (SVM) separation**, aimed at separating the sets \mathcal{A} and \mathcal{B} , on the basis of the following guidelines:

1. Use the polynomial kernel

$$K(x_1, x_2) \triangleq (x_1^T x_2 + c)^d,$$

with $c = 1$ and $d = 2$.

2. Choose the sets \mathcal{A} and \mathcal{B} at your convenience, between the two sets of points (positive with label +1 and negative with label -1).
3. Perform a bilevel 10-fold cross validation, using, for the model selection, a 5-fold cross validation.
4. Compute the following performance indexes:
 - the average training correctness;
 - the average training sensitivity;
 - the average training specificity;
 - the average training F-score;
 - the average testing correctness (accuracy);
 - the average testing sensitivity;
 - the average testing specificity;
 - the average testing F-score.