ESTIN

TP ML (SVM)

EXERCISE 1

Given a binary dataset :

Class0: $\begin{pmatrix} 1 & 9 \\ 5 & 5 \\ 1 & 1 \end{pmatrix}$ **Class1:** $\begin{pmatrix} 8 & 5 \\ 13 & 1 \\ 13 & 9 \end{pmatrix}$

- How many samples (rows) are there (in this dataset)?

- concatenate the two matrices (class0 and class1) and provide the label vector.

- Sketch the points in a scatter plot.

- Sketch the support vectors and the decision boundary for a linear SVM classifier with maximum margin for this data set.

EXERCISE 2

Given a binary dataset:

Class0: $\begin{pmatrix} 2 & 2 \\ 3 & 3 \\ 4 & 4 \\ 5 & 5 \\ 4 & 6 \\ 3 & 7 \\ 4 & 8 \\ 5 & 9 \\ 6 & 10 \end{pmatrix}$ **Class1:** $\begin{pmatrix} 6 & 2 \\ 7 & 3 \\ 8 & 4 \\ 9 & 5 \\ 8 & 6 \\ 7 & 7 \\ 7 & 8 \\ 7 & 9 \\ 8 & 10 \end{pmatrix}$

-Sketch the points in a scatter plot.

- Indicate the support vectors and the decision boundary on the plot you would get using a SVM with linear kernel and a high cost of misclassifying training data (use the SVM module from the sklearn library). What is the error rate of the linear SVM on the training data set?