Machine Learning Models 2024-2025

Exercises Sheet #3

Exercise 1

Given a binary data set:

$$\text{Class} - 1: \begin{bmatrix} 1 & 9 \\ 5 & 5 \\ 1 & 1 \end{bmatrix} \quad \text{Class} + 1: \begin{bmatrix} 8 & 5 \\ 13 & 1 \\ 13 & 9 \end{bmatrix}$$

Plot the points in a plot. Sketch the support vectors and the decision boundary for a linear SVM classifier with maximum margin for this data set.

Exercise 2

Given the binary classification problem:

$$\text{Class -1: } \begin{bmatrix} 2 & 2 \\ 3 & 3 \\ 4 & 4 \\ 5 & 5 \\ 4 & 6 \\ 3 & 7 \\ 4 & 8 \\ 5 & 9 \\ 6 & 10 \end{bmatrix} \quad \text{Class } + 1: \begin{bmatrix} 6 & 2 \\ 7 & 3 \\ 8 & 4 \\ 9 & 5 \\ 8 & 6 \\ 7 & 7 \\ 7 & 8 \\ 7 & 9 \\ 8 & 10 \end{bmatrix}$$

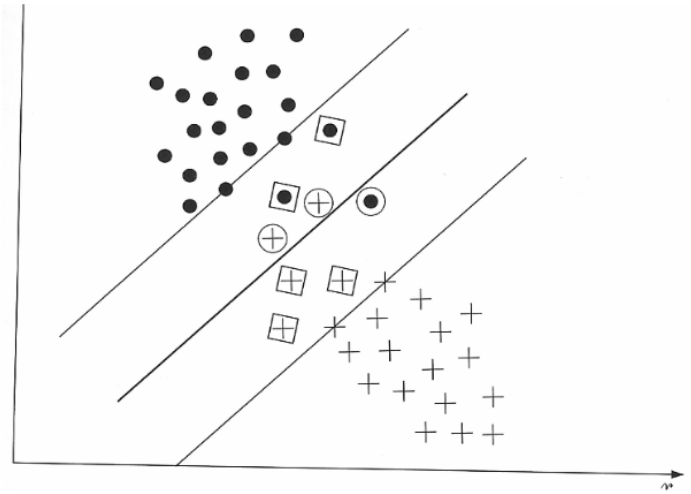
a) Sketch the point in a scatterplot.
b) Sketch on the plot the decision boundary you would get using a SVM with linear kernel and a high cost of misclassifying training data. Indicate the support vectors and the decision boundary on the plot.
c) What is the error rate of the linear SVM on the training data set?

Exercise 3

a) Consider a linear SVM with decision boundary $g(x) = w^T x + w_0$.
In SVM classification, explain why it is useful to assign class labels -1 and 1 for a binary classification problem.

b) The basic SVM optimization problem is to minimize $J = \frac{1}{2} ||w||^2$
What are the additional constraints for this optimization problem?
Ideally, you should answer both by math and explain what this expression means.

c) Explain how slack variables ξ_i are used to solve a non-separable case like the one below:



d) Discuss how likely a Gaussian classifier and an SVM classifier are to overfit to the training data.

e) Explain how an SVM can be used on a classification problem with M classes.