Assignment 2: Support Vector Machines
Machine Learning

Deadline: Friday 01 Nov 2019, 21:00

Introduction

In this assignment, you will further deepen your understanding of Support Vector Machines (SVMs). Please provide a latex based report in the PDF format.

Your report must be archived in a file named "firstname.lastname" and uploaded to the iCorsi website before the deadline expires. Your report should also contain your name. The maximum length of the report must be less than 4 pages. Submissions violating this length constraint might not be evaluated.

Where to get help

We encourage you to use the tutorials to ask questions or to discuss exercises with other students. However, do not look at any report written by others or share your report with others. Violation of that rule will result in 0 points for all students involved.

Grading

The assignment consists of four parts totalling at 100 points. Late submissions (even if a minute late), hand written answers, handing in .doc/docx files, sharing solutions, plagiarism will result in 0 points.

Exercise 1: Kernels I (20 points)

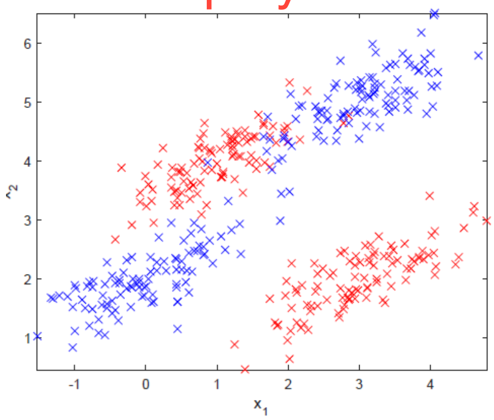
Consider $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d, d \in \mathbb{N}$. Explain in detail why each of the following functions is or is not a valid kernel:

- (6 points) $K(x, y) = x^T y + (x^T y)^2$
- (7 points) $K(x, y) = x^2 e^{-y}$, d = 1
- (7 points) $K(x, y) = ck_1(x, y) + k_2(x, y)$, where $k_1(x, y), k_2(x, y)$ are valid kernels in $\mathbb{R}^d$

Exercise 2: Kernels II (20 points)

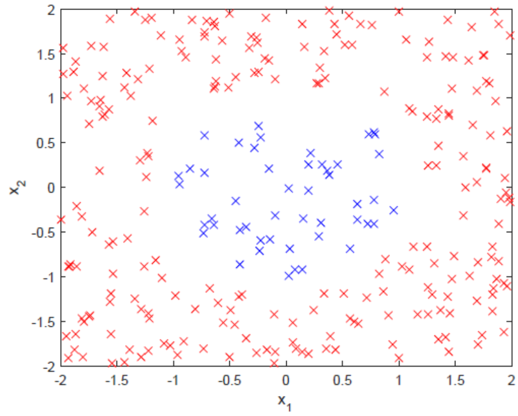
You are provided the following 4 dataset for solving a binary classification problem. You want to preprocess your data before applying a linear solver. Explain in each case if you would apply the kernel trick to represent your data. Which kind of kernel would you use? Why? If relying on kernels is not a good choice, which other transformation would you use?

rbf/polynomial



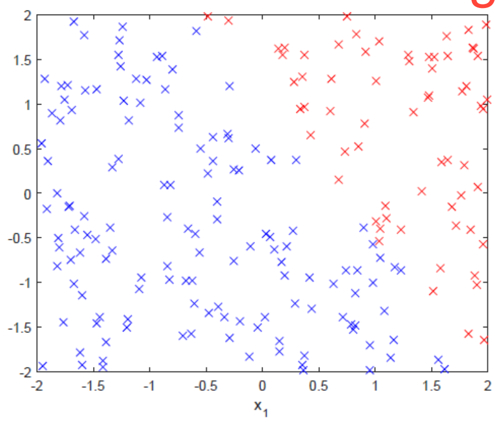
(a)

Polynomial/quadratic/rbf

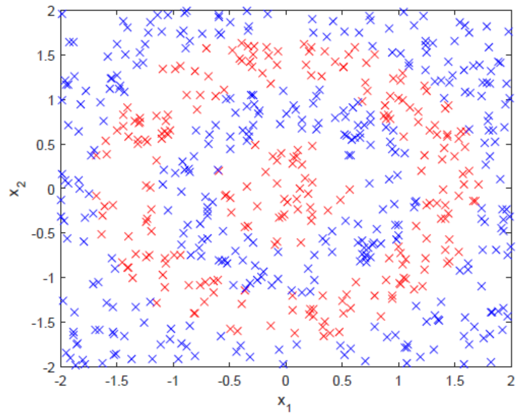


(b)

Linear SVM + softmargin



(c)



(d)

Exercise 3: SVMs (36 points)

Answer the following questions:

- (6 points) Suppose you have 2D examples: is the decision boundary of an SVM with linear kernel a straight line?

Yes

- (6 points) Suppose that the input data are linearly separable. Will an SVM with linear kernel return the same parameters w regardless of the chosen regularization value C? Why?
- (6 points) Suppose you have 3D input examples ($\mathbf{x_i} \in \mathbf{R}^3$). What is the dimension of the decision boundary of the SVM with linear kernel? R2
- (6 points) Is the computational effort for solving a kernel SVM increasing as the dimension of the basis functions increases? Why? No
- (6 points) What is the maximum value of the Gaussian Kernel?
- (6 points) If data are noisy and no linear boundary can perfectly classify all the training data, this means we need to use a feature expansion. True or false?

Exercise 4: SVMs (24 points)

Consider a linear two-class SVM classifier defined by the parameters w = [3; 2] and b = 2. Answer the following questions providing adequate motivations:

- (8 points) Is the point $x_1 = [3; -1]$ classified according to the trained SVM as positive?
- (8 points) Assume to collect a new sample $x_2 = [-2; 2]$ with positive label. Do you need to retrain the SVM?
- (8 points) Does the answer to the previous question apply to a generic new sample?