Exercise for Machine Learning (SS 20)

Assignment 6: Support Vector Machine

Prof. Dr. Steffen Staab, steffen.staab@ipvs.uni-stuttgart.de
Alex Baier, alex.baier@ipvs.uni-stuttgart.de
Janik Hager, janik-manel.hager@ipvs.uni-stuttgart.de
Ramin Hedeshy, ramin.hedeshy@ipvs.uni-stuttgart.de
Analytic Computing, IPVS, University of Stuttgart

Submit your solution in Ilias as either PDF for theory assignments or Jupyter notebook for practical assignments.
Mention the names of all group members and their immatriculation numbers in the file.
Submission is possible until the following Monday, 15.06.2020, at 14:00.

1 Concepts

Explain the following terms and how they are related to SVM in your own words and with (visual) examples:

1. Linear separability
2. Slack variables
3. Kernel functions

2 Perceptron

1. Define the classification function for the perceptron classifier.
2. The dataset for the OR function is given by:

$$X = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$$

$$y = \begin{bmatrix} -1 & 1 & 1 & 1 \end{bmatrix}^\top$$

Given the initial weights of $w = \begin{bmatrix} 1 & -1 & 0.5 \end{bmatrix}$, where w_3 is the bias. Perform the perceptron algorithm (slide 10) with $\alpha = 0.6$ until all data points are correctly classified. Show your computations for each training step. (Note: In the case of $w \cdot x = 0$ output 1.)

3. Prove that the XOR function cannot be represented by a (linear) perceptron.

3 Polynomial Kernel

The second-order polynomial kernel for a two-dimensional vector $x_i = \begin{bmatrix} x_{i1} & x_{i2} \end{bmatrix}^\top$ is defined as:

$$\phi(x_i) = \begin{bmatrix} x_{i1}^2 \\ \sqrt{2} x_{i1} x_{i2} \\ x_{i2}^2 \end{bmatrix}$$

Show that the mapping of the two-dimensional vector to three dimensions is not necessary for calculating the scalar product $\langle \phi(x_i), \phi(x_j) \rangle$. (Note: Transform the equation such that it only uses the scalar product of two-dimensional vectors.)

4 Gaussian Kernel

For all students other than B.Sc. Data Science.
Slide 69 mentions that the Gaussian kernel, also called Radial Basis Function (RBF), projects to an infinite dimensional feature space. Give an intuition on why this is the case and prove it. (Note: Use the Taylor expansion over e^x to show that the Gaussian kernel is an infinite sum over the polynomial kernels.)