University of Stuttgart | ANALYTIC COMPUTING

Foundations of Machine Learning - Exercise (SS 25)

Assignment 9: Support Vector Machine

Arvindh Arunbabu
arvindh.arun@ki.uni-stuttgart.de

Nadeen Fathallah
Nadeen.Fathallah@ki.uni-stuttgart.de

Cosimo Gregucci
cosimo.gregucci@ki.uni-stuttgart.de

Akram Sadat Hosseini
Akram.Hosseini@ki.uni-stuttgart.de

Farane Jalali
farane.jalali-farahani@ki.uni-stuttgart.de

Osama Mohammed
osama.mohammed@ki.uni-stuttgart.de

Jiaxin Pan
jiaxin.pan@ki.uni-stuttgart.de

Tim Schneider
tim.schneider@ki.uni-stuttgart.de

Jingcheng Wu
jingcheng.wu@ki.uni-stuttgart.de

Daniel Frank
daniel.frank@ki.uni-stuttgart.de

Submit your theoretical solution in ILIAS as a single PDF file.[1] Make sure to list the full names of all participants, matriculation number, study program, and B.Sc. or M.Sc on the first page. Optionally, you can *additionally* upload source files (e.g., PPTX files). Submit your programming task in ILIAS as a single Jupyter notebook. If you have any questions, feel free to ask them in the exercise forum in ILIAS.

Submission is open until Monday, 30th of June, 12:00 noon.

[1] Your drawing software probably allows exporting as PDF. An alternative option is to use a PDF printer. If you create multiple PDF files, use a merging tool (like pdfarranger) to combine the PDFs into a single file.

Task 1: SVM

1. **Task** You are training an SVM on a tiny dataset with 4 points shown in Figure 1. This dataset consists of two examples with class label $+1$ (denoted with plus) and two examples with class label -1 (denoted with triangles). By providing **numerical explanations**, answer the following:
   - What is the weight vector w and bias b?
   - What is the equation corresponding to the decision boundary?

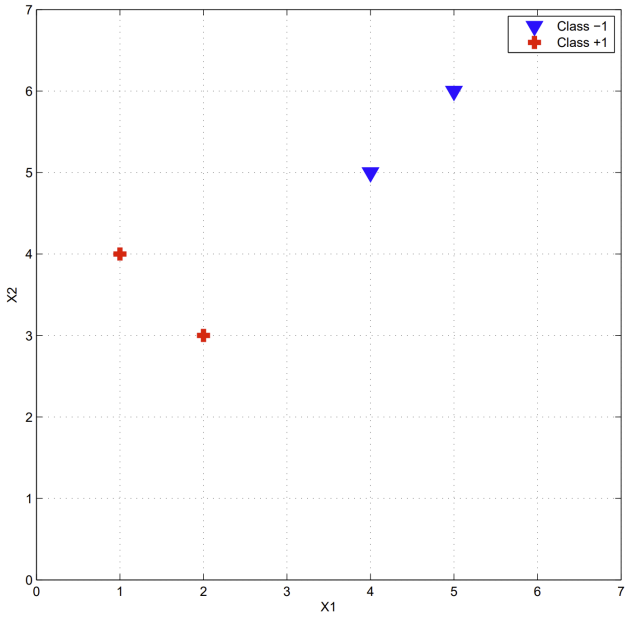

Figure 1 A tiny dataset with 4 points for the SVM task (Task 6.1).

2. **Task** The optimization problem for finding the optimal hyperplane in an SVM for linearly separable data is formulated as:

$$\min_{\mathbf{w},w_0} \frac{1}{2}\|\mathbf{w}\|^2 \tag{1}$$

subject to the constraints:

$$y_i \left(\mathbf{w}^T \mathbf{x}_i + w_0\right) \geq 1, \quad \forall i. \tag{2}$$

Explain why only the support vectors contribute to determining the boundary in the optimization problem.

3. **Task** What are the differences between SVM and Logistic Regression regarding the following aspects:
   - Loss function
   - Objective

Task 2: Kernel Trick

Consider the Radial Basis Function (RBF) kernel defined as:

$$K(x, z) = \exp\left(-\gamma \|x - z\|^2\right) \tag{3}$$

It is known that any valid kernel function $K(x, z)$ corresponds to an implicit feature map $\phi(x)$ such that:

$$K(x, z) = \langle \phi(x), \phi(z) \rangle \tag{4}$$

For example, the polynomial kernel

$$K(x, z) = (\langle x, z \rangle)^2 \tag{5}$$

corresponds to a feature map from $\mathbb{R}^2 \to \mathbb{R}^4$.

However, it is said that the RBF kernel corresponds to a feature map into an **infinite-dimensional** space.

Provide a mathematical explanation for why the RBF kernel implies a feature map into an infinite-dimensional space.

Recall: The Taylor series expansion of the exponential function is given by:

$$e^t = \sum_{n=0}^{\infty} \frac{t^n}{n!} \tag{6}$$

Task 3: Linear SVM

Follow the instructions of **Task 3** in the `09_svm.ipynb` notebook and add your implementation below the lines that are tagged with "`# TODO: ...`". Make sure to have the helper function `svm_helper.py` in the same directory.

Task 4: Kernel Trick - Programming

Follow the instructions of **Task 4** in the `09_svm.ipynb` notebook and add your implementation below the lines that are tagged with "`# TODO: ...`". Make sure to have the helper function `svm_helper.py` in the same directory.