Deadline: Th. Jan. 10, 10:00 am Drop your printed or legible handwritten submissions into the boxes at Samelsonplatz, or upload them as `.pdf` or `.ipynb` files onto the LearnWeb.

Exercise 1 (Separating Hyperplanes - 10 points)**.**

1. (2) Explain what it means for two datasets $X_+, X_- \subset \mathbb{R}^n$ to be *linearly separable*.
2. (4) We have seen in previous exercises that the `XOR`-function is not linearly separable.

x_1	x_2	y
-1	-1	-1
-1	+1	+1
+1	-1	+1
+1	+1	-1

Table 1: `XOR` ($\pm$1-variant)

   Find a function ϕ such that adding $x_3 = \phi(x_1, x_2)$ to the data makes it linearly separable. (provide ϕ and a separating hyperplane!)

3. (4) Find a maximum margin separating hyperplane for the following data:

x_1	x_2	y
0	0	0
1	0.5	1
2	2	0

Exercise 2 (SVMs - 10 points)**.**

1. (2) Explain what the 'kernel trick' is and its purpose.
2. (2) Why is it useful to solve the dual problem, instead of the primal one?
3. (6) In the lecture it was mentioned that both the Primal and Dual SVM problem are Quadratic Programming (QP) problems. Show that this is indeed the case by explicitly construction the matrices A, B, C and vectors a, b, c.

SVM Primal	SVM Dual	QP Problem
$\min\limits_{\beta,\xi} \frac{1}{2}\|\beta\|^2 + \gamma \sum_{i=1}^{n} \xi_i$	$\max\limits_{\alpha} -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j \langle x_i \vert x_j \rangle + \sum_{i=1}^{n} \alpha_i$	$\min\limits_{z} \frac{1}{2} z^\intercal C z + c^\intercal z$
$y_i(\beta_0 + \langle \beta \vert x_i \rangle) \geq 1 - \xi_i$	$\sum_{i=1}^{n} \alpha_i y_i = 0$	$Az = a$
$\xi_i \geq 0$	$0 \leq \alpha_i \leq \gamma$	$Bz \leq b$

For example, with $z = \binom{\beta}{\xi}$, the objective function of the primal can be rewritten as:

$$\min_{\beta,\xi} \frac{1}{2}\|\beta\|^2 + \gamma \sum_{i=1}^{n} \xi_i \iff \min_z \frac{1}{2} z^T \underbrace{\left[\begin{array}{c|c} I & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right]}_{=C} z + \underbrace{\left[0 \cdots 0 | \gamma \cdots \gamma\right]}_{=c} \cdot z$$