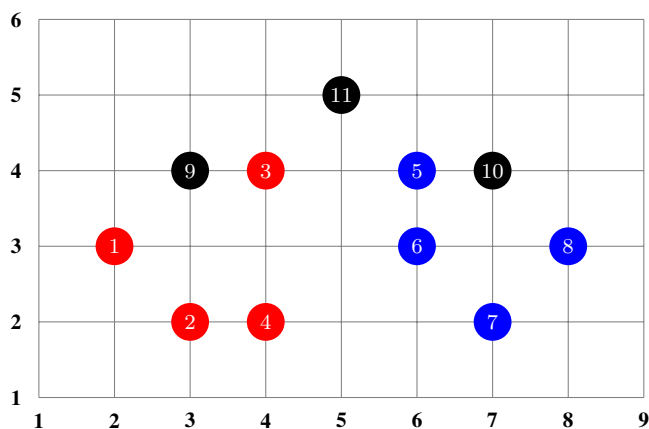


Knowledge Discovery and Data Mining I
 WS 2019/20

Exercise 4: SVM, Kernel Trick, Linear Separability

Exercise 4-1 Support Vector Machines



Consider the following training data:

$$x_1 = (2, 3), x_2 = (3, 2), x_3 = (4, 4), x_4 = (4, 2)$$

$$x_5 = (6, 4), x_6 = (6, 3), x_7 = (7, 2), x_8 = (8, 3)$$

Let $y_A = -1, y_B = +1$ be the class indicators for both classes

$$A = \{x_1, x_2, x_3, x_4\}, B = \{x_5, x_6, x_7, x_8\}.$$

- (a) Just using the above-standing plot, specify which of the points should be identified as support vectors.
- (b) Draw the maximum margin line which separates the classes (you don't have to do any computations here). Write down the normalized normal vector $w \in \mathbb{R}^2$ of the separating line and the offset parameter $b \in \mathbb{R}$.
- (c) Consider the decision rule: $H(x) = \langle w, x \rangle + b$. Explain how this equation classifies points on either side of a line. Determine the class for the points $x_9 = (3, 4), x_{10} = (7, 4)$ and $x_{11} = (5, 5)$.

Exercise 4-2 Kernel Trick

Consider the polynomial kernel function

$$K : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto (x^T y + \gamma)^p, \text{ with } p = 2, \gamma = 1.$$

Furthermore let

$$\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^6, x \mapsto (1, \sqrt{2}x_1, \sqrt{2}x_2, x_1^2, x_2^2, \sqrt{2}x_1x_2).$$

Show that $K(x, y) = \langle \phi(x), \phi(y) \rangle$.

Exercise 4-3 Mercer Kernels

As known from the lecture, a Mercer kernel $\kappa : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ needs to fulfil

- (1) Symmetry, i.e., $\kappa(x, y) = \kappa(y, x)$
- (2) Positive semi-definiteness, i.e. the kernel matrix $\kappa(X) := (\kappa(x_i, x_j))_{ij} \in \mathbb{R}^n$ is positive semi-definite for all $X = \{x_1, \dots, x_n\} \subseteq \mathcal{X}$.

Show that the following functions are Mercer kernels for $x, y \in \mathcal{X} = \mathbb{R}^d$.

- (a) $\kappa_1(x, y) = \begin{cases} 1 & x = y \\ 0 & x \neq y \end{cases}$
- (b) $\kappa_2(x, y) = x^T y$.
- (c) $\kappa_3(x, y) = \alpha x^T y + \beta$ for $\alpha, \beta \in \mathbb{R}$ with $\alpha, \beta \geq 0$

Exercise 4-4 Linear Separability

In the following exercise, provide minimal subsets $\{x_1, \dots, x_m\} = X \subseteq \mathcal{X} = \mathbb{R}^d$ together with class labels $y_1, \dots, y_m \in \{-1, 1\}$ for the given dimensionality $d \in \mathbb{N}$ that are not linear separable. Prove both, the minimality (i.e. every $X' \subseteq X$ with $|X'| < |X|$ is linearly separable), as well as the non-separability of X .

- (a) $d = 1$
- (b) $d = 2$