Neural Networks and Learning Theory

Prof. Dr. Klaus Meer, Ameen Naif

Exercise Sheet 4
Version 14. Juni 2017

Exercise 1.
Consider the case of a hyperplane for linearly separable patterns, wich is defined by the equation:

$$w^T x + b = 0$$

where w denotes the weight vector, b denotes the bias, and x denotes the input point.

(a) Define the optimal hyperplane in this case (linearly separable patterns).

(b) Find the distance between the optimal hyperplane and the closest input point.

(c) Show that the optimal hyperplane is unique.

Exercise 2.
Consider the nonlinear transformation mapping $\varphi : \mathbb{R} \to \mathbb{R}^{d+1}$ from the input space $\mathbb{R}$ to the feature space $\mathbb{R}^{d+1}$ defined by

$$\varphi(x) = (a_0 \cdot 1, a_1 \cdot x, a_2 \cdot x^2, \ldots, a_d \cdot x^d), a_i \in \mathbb{R} \ .$$

We introduce the **inner-product kernel** $k(x, y)$ defined by $k(x, y) := \varphi(x)^T \cdot \varphi(y)$.

a) Detrmine K in the case, where alle a_i's equal 1.

b) Find the values a_i such that $k(x, y) = (1 + x \cdot y)^d$. Compare your result with the inner-product kernel in a).

c) In case a) show the following: for every choice of $d+1$ different points $x_i, 0 \le i \le d$ and for any target output $d_i \in \{-1, 1\}, 0 \le i \le d$ there is a unique $w \in \mathbb{R}^{d+1}$ with $w^T \cdot \varphi(x_i) = d_i$ for all i.

Exercise 3.
Consider the inner-product kernel $k(x, y)$ defined by

$$k(x, y) = (1 + x \cdot y)^d$$

where $x, y \in \mathbb{R}^n$ and $d \in \mathbb{N}$ fixed.

a) Show that there exists an integer N and a transformation mapping $\varphi : \mathbb{R} \to \mathbb{R}^N$ from the input space $\mathbb{R}^n$ to the feature space $\mathbb{R}^N$, s.t. $k(x, y) = \varphi(x)^T \cdot \varphi(y)$.

b) In case a) find the mapping φ for $d := 3$ and $n := 2$.

c) What is the minimum value of the power d for which the XOR problem can be solved? What is the result of using a value for d larger than the minimum?