Exercise Sheet 7

Notation

1. Let $\langle x, x' \rangle$ denote the scalar product of points $x, x' \in \mathbb{R}^d$.

2. The sign function is defined as

$$\operatorname{sgn} : \mathbb{R} \to \mathbb{R}, \quad u \mapsto \begin{cases} +1 & : \quad u \geq 0 \\ -1 & : \quad u < 0 \end{cases}.$$

Exercise 1

The hypothesis class of homogeneous halfspaces (linear classifiers without bias) in $\mathbb{R}^d$ is defined by

$$\mathcal{H} = \{f : \mathbb{R}^d \to \{\pm 1\} \;:\; \exists\, w \in \mathbb{R}^d \text{ s.t. } f(x) = \operatorname{sgn}\big(\langle w, x \rangle\big)\}.$$

Show that the VC dimension of $\mathcal{H}$ is d.

Hints: First show (a) that $\mathcal{H}$ shatters d points in $\mathbb{R}^d$. Next show (b) that $\mathcal{H}$ cannot shatter any set of $d+1$ points in $\mathbb{R}^d$. Exploit linear independence in (a) and linear dependence in (b).

Exercise 2

The hypothesis class of inhomogeneous halfspaces (linear classifiers) in $\mathbb{R}^d$ is defined by

$$\mathcal{H} = \{f : \mathbb{R}^d \to \{\pm 1\} \;:\; \exists\, w \in \mathbb{R}^d, \exists\, b \in \mathbb{R} \;\text{ s.t. }\; f(x) = \operatorname{sgn}\big(\langle w, x \rangle + b\big)\}.$$

Show that the VC dimension of $\mathcal{H}$ is $d+1$.

Exercise 3

Let $\mathcal{X}$ be a space endowed with a metric $\delta : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$. What is the VC dimension of the 1-nearest-neighbor classifier in $(\mathcal{X}, \delta)$? Discuss the implications of your result.

Note: You can think of $(\mathcal{X}, \delta)$ as the Euclidean space $(\mathbb{R}^d, \|\cdot\|_2)$.

Exercise 4

Plot the VC bound as a function of the sample size n and as function of the error tolerance ε for VC dimensions 1, 3 and 5.