Thomas Kesselheim
Alexander Braun

21. April 2022

Algorithmische Grundlagen des Maschinellen Lernens

Sommersemester 2022

Präsenzblatt 3

Aufgabe 1:

Betrachten Sie für ein $p \in [0,1]$ voneinander unabhängige Münzwürfe gegeben durch die Zufallsvariablen $Z_1, \ldots, Z_n$. Dabei gilt für alle $1 \leq i \leq n$, dass $\mathbf{Pr}[Z_i = 1] = p$ und $\mathbf{Pr}[Z_i = 0] = 1 - p$. Sei $\bar{Z} = \frac{1}{n}\sum_{i=1}^{n} Z_i$. Zeigen Sie mithilfe der Hoeffding-Ungleichung, dass

$$\mathbf{Pr}\left[|\bar{Z} - p| \geq \frac{1}{4}\right] \leq 2\exp(-\frac{n}{8}).$$

Aufgabe 2:

Begründen Sie jeweils kurz, ob folgende Aussagen richtig oder falsch sind. Im Folgenden bezeichnet $\mathcal{H}$ eine Hypothesenklasse und $\mathcal{R}$ ein Mengensystem über einer Grundmenge $\mathcal{X}$.

(a) Hat $\mathcal{H}$ endliche VC-Dimension, so ist für $S \subset \mathcal{X}$ die Hypothesenklasse $\mathcal{H}|_S$ PAC-lernbar.

(b) Besteht $\mathcal{H}$ aus Funktionen der Form $h_a \colon \mathbb{R} \to \{+1, -1\}$ mit $h_a(x) = +1$ genau dann, wenn $x \geq a$ für $a \in \mathbb{R}$, so hat $\mathcal{H}$ VC-Dimension 2.

(c) Ist $\dim(\mathcal{R}) = m$, so gilt $\Pi_{\mathcal{R}}(m) = 2^m$.

(d) Ist $\Pi_{\mathcal{R}}(m) = 2^m$, so gilt $\dim(\mathcal{R}) = m$.