Hans U. Simon
Malte Darnstädt

Bochum, den 02.05.2014
Abgabe bis zum 12.05.2014

Übungen zur Vorlesung

Theorie des Maschinellen Lernens

Sommer 2014

Übungsblatt 04

Aufgabe 4.1

Sei eine Wahrscheinlichkeitsverteilung über einer Menge X gegeben. Sei $A \subseteq X$ mit $\Pr(A) = \epsilon$. Zeige, dass eine Stichprobe der Größe

$$m = \Omega\left(\frac{k}{\epsilon}\right)$$

notwendig ist, damit mindestens k Punkte aus A gezogen werden (mit einer Wahrscheinlichkeit $\geq 1/2$).

Bemerkung: In Präsenzaufgabe 1.4 wurde gezeigt, dass $m = O\left(\frac{k+\log(1/\delta)}{\epsilon}\right)$ viele Beispiele ausreichen um das gewünschte Ziel mit Wahrscheinlichkeit $\geq 1-\delta$ zu erreichen.

Aufgabe 4.2

Sei $X_n = \mathbb{R}$. Wir betrachten die Konzeptklasse P_k, die durch das Vorzeichen von Polynomen von Grad k gebildet wird; das heißt die Elemente von P_k haben die folgende Form:

$$c(x) = \begin{cases} 1 & \text{falls } \sum_{i=0}^{k} a_i x^i \geq 0 \\ 0 & \text{sonst} \end{cases}$$

mit $a_i \in \mathbb{R}$.

Bestimme die VC-Dimension von P_k.

Aufgabe 4.3

Sei X eine Menge und $\mathcal{H}_1, \mathcal{H}_2$ seien zwei Hypothesenklassen über X. Bestimme für die folgenden Ausdrücke möglichst genaue untere und obere Schranken in Abhängigkeit von $\Pi_{\mathcal{H}_1}(m)$ und $\Pi_{\mathcal{H}_2}(m)$ oder von Konstanten. Zeige, dass auch im Fall $\mathcal{H}_1 \neq \mathcal{H}_2$ in jeder Ungleichung Gleichheit gelten kann.

a) $\Pi_{\mathcal{H}_1 \cap \mathcal{H}_2}(m)$　　b) $\Pi_{\mathcal{H}_1 \cup \mathcal{H}_2}(m)$　　c) $\Pi_{\mathcal{H}}(m)$
wobei $\mathcal{H} = \{h_1 \cup h_2 \mid h_1 \in \mathcal{H}_1, h_2 \in \mathcal{H}_2\}$

Aufgabe 4.4

Zeige, dass die Schranke in Sauers Lemma scharf ist: Für jedes d existiert eine Klasse $\mathcal{H}$ mit VC-Dimension d, so dass für alle $m \in \mathbb{N}$ gilt

$$\Pi_{\mathcal{H}}(m) = \sum_{i=0}^{d} \binom{m}{i}$$