Hans U. Simon
Christoph Ries

Bochum, den 24. April 2018
Abgabe bis zum 02.05.2018

Übungen zur Vorlesung

Theorie des maschinellen Lernens

Sommer 18

Übungsblatt 02

Aufgabe 2.1 (4 Punkte)

Wir betrachten die Klasse aller achsenparallelen Rechtecke

$$\mathcal{H}_{rec}^2 := \{h_{(a_1,b_1,a_2,b_2)} : a_1 \leq b_1 \wedge a_2 \leq b_2\}.$$

Eine Hypothese aus $\mathcal{H}_{rec}^2$ hat die Form

$$h_{(a_1,b_1,a_2,b_2)}(x_1,x_2) = \begin{cases} 1 & \text{, falls } a_1 \leq x_1 \leq b_1 \wedge a_2 \leq x_2 \leq b_2 \\ 0 & \text{, sonst} \end{cases}.$$

Wir betrachten im Folgenden den realisierbaren Fall.

a) Sei A ein Algorithmus. A bestimme das kleinste Rechteck, welches alle positiven Punkte der Trainingsmenge umschließt, und gebe die zugehörige Hypothese aus. Zeige, dass A ein ERM-Algorithmus ist.

b) Skizziere analog zu den Halbintervallen aus der Vorlesung einen Beweis, dass A bei Eingabe einer Trainingsmenge S mit

$$|S| \geq \frac{4\log(4/\delta)}{\varepsilon}$$

mit einer Wahrscheinlichkeit von mindestens $1-\delta$ eine Hypothese ausgibt deren Fehler höchstens ε ist.

c) Wieviele Beispiele braucht ein entsprechender Algorithmus für achsenparallele d-dimensionale Boxen (ein Rechteck ist eine zwei-dimensionale Box) in $\mathbb{R}^d$? Beweise deine Behauptung.

Aufgabe 2.2 (4 Punkte)

Sei $\mathcal{H}$ eine Hypothesenklasse (ggf. unendlich), die die uniforme Konvergenzbedingung bezüglich einer Verlustfunktion l erfüllt. Zeige, dass $\mathcal{H}$ agnostisch PAC-lernbar bezüglich l ist und dass gilt

$$m_{\mathcal{H}}(\varepsilon,\delta) \leq m_{\mathcal{H}}^{UC}(\varepsilon/3,\delta).$$

— Bitte wenden! —

Aufgabe 2.3 (4 Punkte)
Sei $\mathcal{H}_k$ die Klasse aller konvexen Polygonen mit bis zu k Ecken.
Zeige:

$$\mathrm{VCdim}(\mathcal{H}_k) = 2k + 1$$

Aufgabe 2.4 (4 Punkte)
Sei $X = \mathbb{R}$. Die Hypothesenklasse sei durch

$$\mathcal{H} := \{x \mapsto \lceil 0.5 \cdot \sin(\theta x) \rceil \,:\, \theta \in \mathbb{R}\}$$

gegeben. Zeige, dass $\mathrm{VCdim}(\mathcal{H}) = \infty$.

TIPP: *Eine Möglichkeit besteht darin, folgende Aussage mit Beweis zu verwenden. Sei $0.x_1x_2x_3...$ die binäre Darstellung von $x \in (0, 1)$, dann gilt für alle $m \in \mathbb{N}$*

$$\lceil 0.5 \cdot \sin(2^m \pi x) \rceil = 1 - x_m$$

vorausgesetzt, dass ein $k \geq m$ existiert mit $x_k = 1$.