Hans U. Simon
Christoph Ries

Bochum, den 3. Mai 2016
Abgabe bis zum 10.05.2016

Übungen zur Vorlesung

Theorie des maschinellen Lernens

Sommer 16

Übungsblatt 03

Hinweis: Die für dieses Blatt nötigen Concentration Inequalities sind auf der Homepage

`http://www.ruhr-uni-bochum.de/lmi/lehre/ml_ss16/`

unter dem Unterpunkt Materialien zu finden.

Aufgabe 3.1 (4 Punkte)
Zeige, dass die Schranke in dem Lemma von Sauer scharf ist, d.h. für alle $d \in \mathbb{N}$ existiert eine Klasse C, sodass für alle $m \in \mathbb{N}$ gilt

$$\Pi_C(m) = \sum_{i=0}^{d} \binom{m}{i}.$$

Aufgabe 3.2 (4 Punkte)
Wir wollen die Ungleichung

$$Pr_{S\sim D^m}\{|L_D(h) - L_S(h)| \le \varepsilon\} \ge \frac{1}{2}. \qquad (\star)$$

für eine feste Hypothesen $h : \mathcal{X} \to \{0,1\}$ mit Hilfe der Chebyshev-Ungleichung zeigen. Gehe dazu wie folgt vor.

a) Zunächst müssen die Symbole aus der Chebyshev-Ungleichung, wie sie in der Version auf der Homepage auftauchen, belegt werden. Wie sieht die Zufallsvariable X aus? Was ist der Erwartungswert und die Varianz von X? Wie muss t gewählt werden?

b) Zeige, dass obige Ungleichung $(\star)$ durch die Chebyshev-Ungleichung folgt, sofern $m \ge \frac{1}{2\varepsilon^2}$.

— Bitte wenden! —

Aufgabe 3.3 (4 Punkte)

Seien $\mathcal{X}$ die Grundmenge und D eine Verteilung über $\mathcal{X} \times \{0, 1\}$. $T = (z_1, ..., z_{2m})$ sei eine $2m$-elementige Trainingsmenge, die zufällig bzgl. D^{2m} generiert wird. Jedes Element aus T hat die Form $(x_i, y_i) = z_i$.
$\mathcal{S} := \{\sigma_I : I \subseteq [m]\}$ sei eine Teilmenge aller Permutationen über $\{1, ..., 2m\}$. Eine Permutation $\sigma_I \in \mathcal{S}$ tauscht z_i mit z_{m+i} für alle $i \in I$. Wird $\sigma \in \mathcal{S}$ auf T angewendet, wird dies als $T_\sigma = (z_{\sigma(1)}, z_{\sigma(2)}, ..., z_{\sigma(2m)})$ notiert. Wir zerlegen T_σ in $T_\sigma^1 = (z_{\sigma(1)}, ..., z_{\sigma(m)})$ und $T_\sigma^2 = (z_{\sigma(m+1)}, ..., z_{\sigma(2m)})$.
Sei h nun eine feste Hypothese. Wir wollen die Wahrscheinlichkeit über die zufällige Wahl von $\sigma \in \mathcal{S}$ für das Ereignis

$$|L_{T_\sigma^1}(h) - L_{T_\sigma^2}(h)| \geq \frac{\varepsilon}{2}$$

nach oben abschätzen. Beachte, dass der Wahl von σ eine uniforme Verteilung zugrunde liegt.

a) Zur Abschätzung verwenden wir die Ungleichung von Hoeffding aus den Materialien von der Homepage. Was ist n? Wie müssen die Zufallsvariablen $X_1, ..., X_n$ gewählt werden? Was ist ihr Erwartungswert? Welche Werte sind für a und b einzusetzen?

b) Wende nun die Ungleichung von Hoeffding an. Wie sieht die obere Schranke aus?

Aufgabe 3.4 (4 Punkte)

Zeige, dass folgende Hypothesenklasse nicht-uniform lernbar ist.

$$\mathcal{H} := \{h_p \mid p : \mathbb{R} \to \mathbb{R} \text{ ist ein Polynom}\},$$

wobei $h_p(x) = sign(p(x))$. Ein Punkt $x \in \mathbb{R}$ wird von einer Hypothese $h_p \in \mathcal{H}$ genau dann mit 1 markiert wird, wenn $p(x) \geq 0$.