Hans U. Simon
Malte Darnstädt

Bochum, den 30.04.2013
Abgabe bis zum 08.05.2013

Übungen zur Vorlesung
Theorie des Maschinellen Lernens
Sommer 2013
Übungsblatt 03

In Aufgabe 3.1 und 3.2 untersuchen wir im Restricted Model folgende Klasse:

Sei $X = \mathbb{N}$ und T_d die Klasse der Teilmengen von X mit höchstens d Elementen (wir haben Klassen eigentlich als Mengen von $\{0,1\}$-wertigen Funktionen definiert; hier identifizieren wir die Mengen in T_d mit ihren Indikator-Funktionen).

Aufgabe 3.1 (4 Punkte)

Sei L_1 der Lerner, der die zur Stichprobe z *kleinste* konsistente Menge aus T_d ausgibt. Zeige, dass der Faktor von $\log(1/\epsilon)$ in der oberen Schranke verschwindet. Zeige also:

$$m_{L_1}(\epsilon, \delta) = O\left(\frac{d + \log(1/\delta)}{\epsilon}\right)$$

Aufgabe 3.2 (4 Punkte)

Sei L_2 ein Lerner, der eine zur Stichprobe z *größte* konsistente Menge aus T_d ausgibt. Da es im Allgemeinen mehrere solcher Mengen gibt und wir ein eindeutiges Verhalten von L_2 wünschen, vereinbaren wir, dass L_2 die größte konsistente Menge mit den numerisch kleinstmöglichen Elementen wählt.

Zeige, dass sich der Faktor von $\log(1/\epsilon)$ in der oberen Schranke nicht vermeiden lässt. Zeige also:

$$m_{L_2}(\epsilon, \delta) = \Omega\left(\frac{d\log(1/\epsilon) + \log(1/\delta)}{\epsilon}\right)$$

Aufgabe 3.3 (8 Punkte)

Sei A ein randomisierter Algorithmus, der eine Klasse H mit $m_A(\epsilon, \delta)$ gelabelten Beispielen und $r_A(\epsilon, \delta)$ Würfen einer fairen Münze lernt. Dabei sei in H die konstante Null- und die konstante Eins-Funktion enthalten. Sei $0 < \epsilon \leq \frac{1}{4}$, $0 < \delta < 1$ beliebig.

a) (2 Punkte) Wie kann man mit Hilfe von gelabelten Beispielen eine faire Münze simulieren?

b) (2 Punkte) Angenommen es gelte $P\{(x,y) : y = b\} < \epsilon$ für ein $b \in \{0,1\}$. Mit welcher Strategie kann man in dieser Situation lernen? Wie viele Beispiele reichen zum Erfüllen des ϵ-δ-Kriteriums aus?

c) (4 Punkte) Zeige, dass ein deterministischer Algorithmus B existiert, der mit

$$m_A(\epsilon, \delta/2) + O\left(\frac{\log(1/\delta) + r_A(\epsilon, \delta/2)}{\epsilon}\right)$$

gelabelten Beispielen das ϵ-δ-Kriteriums erfüllt (also H erfolgreich lernt).

Hinweis Bei Aufgabe 3.3b) und c) kann man die multiplikative Chernoff-Schranke verwenden:

Seien $x_1, \ldots, x_m$ unabhängige $\{0, 1\}$-wertige Zufallsvariablen. Sei $p_i := \Pr(x_i = 1)$ und $\mu := \frac{1}{m} \sum_{i=1}^{m} p_i$. Dann gilt für $\epsilon \geq 0$:

$$\Pr\left(\frac{1}{m} \sum_{i=1}^{m} x_i \geq (1 + \epsilon)\mu\right) \leq \exp(-\epsilon^2 \mu m/3)$$

und

$$\Pr\left(\frac{1}{m} \sum_{i=1}^{m} x_i \leq (1 - \epsilon)\mu\right) \leq \exp(-\epsilon^2 \mu m/2)$$