Hans U. Simon
Malte Darnstädt

Bochum, den 27.06.2014
Abgabe bis zum 8.07.2014

Übungen zur Vorlesung

Theorie des Maschinellen Lernens

Sommer 2014

Übungsblatt 11

Aufgabe 11.1

Es seien drei Variablen A, B, C mit τ-genauen Approximationen $\hat{A}, \hat{B}, \hat{C}$ gegeben, wobei gilt $0 \leq A, \hat{A} \leq A_1$, $0 < B_0 \leq B, \hat{B} \leq B_1$ und $0 < C_0 \leq C, \hat{C} \leq C_1$.
Wir wollen den Fehler $\left|\frac{\hat{A}}{\hat{B}\hat{C}} - \frac{A}{BC}\right|$ bestimmen.

- Verwende zunächst die Fehlerfortpflanzungsregeln aus der Vorlesung (angepasst für unterschiedliche τ) um den Fehler für die Zerlegung $A \cdot \frac{1}{B \cdot C}$ zu bestimmen.
- Entwickle dann eine neue Fehlerfortpflanzungsregel für das Produkt aus drei Variablen und analysiere die Zerlegung $A \cdot \frac{1}{B} \cdot \frac{1}{C}$ entsprechend.

Welche Analyse ergibt die bessere Fehlerabschätzung?

Aufgabe 11.2

Beweise mit Hilfe der additiven Chernoff-Schranke: Wenn eine Klasse $\mathcal{C}$ durch statistische Anfragen effizient lernbar ist, so ist sie auch effizient PAC-lernbar (im Modell ohne Rauschen).

Hinweis: Ein Spezialfall der additiven Chernoff-Schranke lautet:
Seien $X_1, \ldots, X_m$ unabhängige, identisch verteilte $\{0,1\}$-wertige Zufallsvariablen. Sei $p := \Pr(X_i = 1)$. Dann gilt für $\gamma \geq 0$:

$$\Pr\left(\left|\frac{1}{m}\sum_{i=1}^{m} X_i - p\right| \geq \gamma\right) \leq 2e^{-2\gamma^2 m}$$

Aufgabe 11.3

Es seien zwei gezinkte Münzen gegeben; eine zeigt häufiger Kopf, die andere zeigt mit derselben Wahrscheinlichkeit häufiger Zahl. Eine der Münzen wird zufällig ausgewählt und es werden m unabhängige Würfe durchgeführt.

Es soll ermittelt werden welche Münze ausgewählt wurde.

Wie lautet die beste Entscheidungsregel? Beweise, dass keine Strategie eine höhere Wahrscheinlichkeit besitzt richtig zu entscheiden als deine Regel.

Aufgabe 11.4

Sei $\mathcal{C}$ eine Klasse mit $|\mathcal{C}| > 1$. Zeige, dass ein gegen Klassifikationsrauschen resistenter PAC-Lerner mindestens

$$m = \Omega\left(\frac{1}{(1-2\eta)^2}\right)$$

viele Beispiele benötigt.

Hinweis: Es reicht aus, dies für ein konstantes ϵ und δ zu zeigen (z.B. $\epsilon = \delta = 1/4$) und für alle $1/4 \leq \eta < 1/2$. Verwende deine Strategie von Aufgabe 11.3 und folgendes Resultat aus der Wahrscheinlichkeitstheorie:
Seien $X_1, \ldots, X_m$ unabhängige, identisch verteilte $\{0,1\}$-wertige Zufallsvariablen und $p := \Pr(X_i = 1)$. Dann gilt für alle k mit $mp \leq k \leq m(1-p)$:

$$\Pr\left(\sum_{i=1}^{m} X_i \geq k\right) \geq \frac{1}{2}\left(1 - \sqrt{1 - e^{-\frac{(k-mp)^2}{mp(1-p)}}}\right)$$