

Übungen zur Vorlesung
Theorie des maschinellen Lernens
Sommer 18
Übungsblatt 07

Aufgabe 7.1 (4 Punkte)

- a) Betrachte die Trainingsmenge $S = \{((0, 0), 1), ((1, 0), -1), ((0, 1), -1), ((1, 1), 1)\}$. Lasse AdaBoost mit Entscheidungsstümpfen als Basislerner auf S laufen. Was ist der kleinste Fehler den AdaBoost auf S realisieren kann? Begründe. Beachte, dass der Basislerner in diesem Fall kein γ -schwacher Lerner ist.
- b) Welche Hypothese gibt AdaBoost mit Entscheidungsstümpfen als Basislerner nach zwei Runden auf der Trainingsmenge

$$S = \{(1, 1), (2, 1), (3, 1), (4, -1), (5, -1), (6, -1), (7, 1)\}$$

aus.

Aufgabe 7.2 (4 Punkte)

Betrachte den Algorithmus AdaBoost. Zeige, dass der Fehler von h_t bezüglich der Verteilung $D^{(t+1)}$ exakt $1/2$ ist, d.h. zeige, dass für alle $t \in [T]$

$$\sum_{i=1}^m D_i^{(t+1)} \mathbf{1}_{[y_i \neq h_t(x_i)]} = 1/2.$$

— Bitte wenden! —

Aufgabe 7.3 (4 Punkte)

Sei A ein Lernalgorithmus. Es existiere ein $\delta_0 \in (0, 1)$ und eine Funktion $m_{\mathcal{H}} : (0, 1) \rightarrow \mathbb{N}$, sodass für alle $\varepsilon \in (0, 1)$, alle Verteilungen D und alle $m \geq m_{\mathcal{H}}(\varepsilon)$, gilt

$$\Pr_{S \sim D^m} [L_D(A(S)) \leq \min_{h \in \mathcal{H}} L_D(h) + \varepsilon] \geq 1 - \delta_0.$$

Schlage ein Verfahren vor, welches A verwendet und \mathcal{H} agnostisch PAC-lernt mit

$$m_{\mathcal{H}}(\varepsilon, \delta) \leq k \cdot m_{\mathcal{H}}(\varepsilon) + \left\lceil \frac{2 \log(4k/\delta)}{\varepsilon^2} \right\rceil,$$

wobei $k = \lceil \log(\delta) / \log(\delta_0) \rceil$.

Aufgabe 7.4 (4 Punkte)

Bei der k -fachen Kreuz-Validierung wird eine Trainingsmenge S der Größe m in k Teile S_1, \dots, S_k der Größe m/k zerlegt. Der Fall $k = m$ heißt Leave-one-out (LOO).

Gegeben sei eine Grundmenge \mathcal{X} mit einer Verteilung D . Eine Instanz $x \in \mathcal{X}$ wird zufällig mit $y = 1$ bzw. $y = 0$ markiert - es soll dabei gelten $\Pr[y = 1 \mid x] = \Pr[y = 0 \mid x] = 1/2$. Sei $S = \{(x_1, y_1), \dots, (x_m, y_m)\}$ eine Trainingsmenge. Ein Lernalgorithmus A liefere bei Eingabe von S die Hypothese h , wobei

$$h(x) = \begin{cases} 1 & , \text{ falls } \sum_{i=1}^{|S|} y_i = 1 \pmod{2} \\ 0 & , \text{ sonst} \end{cases}.$$

Gemäß LOO sei $S_i = \{(x_i, y_i)\}$ für alle $1 \leq i \leq k = m$. Zeige, dass der mittlere Fehler

$$\frac{1}{k} \sum_{i=1}^k L_{S_i}(A(S \setminus S_i))$$

entweder 0 oder 1 ist, obwohl der tatsächliche Fehler (bei komplett zufälligen Labels) $1/2$ ist.