

Algorithmische Grundlagen des Maschinellen Lernens

Sommersemester 2022

Übungsblatt 3

Aufgabe 1: (3+3 Punkte)

Sei \mathcal{H} eine Hypothesenklasse mit Grundmenge \mathcal{X} . Das heißt, \mathcal{H} ist eine Menge von Funktionen der Form $h: \mathcal{X} \rightarrow \{-1, +1\}$. Zeigen Sie, dass eine endliche Grundmenge \mathcal{X} die PAC-Lernbarkeit von \mathcal{H} impliziert.

- (a) Betrachten Sie den realisierbaren Fall.
- (b) Betrachten Sie den nicht-realisierbaren Fall.

Aufgabe 2: (3+3 Punkte)

Wir betrachten wieder die Hypothesenklasse \mathcal{H} der Schwellenwertfunktionen. Betrachten Sie die folgenden Verteilungen \mathcal{D} über Paare (x, y) , wobei $x \in \mathbb{R}$ ein Datenpunkt ist und $y \in \{-1, +1\}$ das zugehörige Label. Bestimmen Sie jeweils ein Hypothese $h \in \mathcal{H}$, die den tatsächlichen Fehler $\text{err}_{\mathcal{D}}(h)$ minimiert, und beweisen Sie Ihre Behauptung.

- (a) Betrachte die Dichtefunktion $g(x) = e^{-x}$ für $x \geq 0$ und $g(x) = 0$ für $x < 0$. Ziehe x aus g und setze $y = 1$, falls $x \in [a, b]$, für $a > 0$ und $b > a + 1$, sonst $y = -1$.
- (b) Ziehe x uniform zufällig aus $[-2, 2]$. Bestimme anschließend y wie folgt: Wenn $x < 0$, setze $y = -1$; wenn $x > 1$, setze $y = +1$; anderenfalls setze $y = +1$ mit Wahrscheinlichkeit x^2 und $y = -1$ mit Wahrscheinlichkeit $1 - x^2$.

Aufgabe 3: (4+4 Punkte)

Sei \mathcal{H} die Menge aller Funktionen $h: \mathbb{R} \rightarrow \{-1, +1\}$ der Form

(a)
$$h_a(x) = \begin{cases} +1 & \text{falls } x \geq a \\ -1 & \text{sonst} \end{cases}$$

(b)
$$h_{a,b}(x) = \begin{cases} +1 & \text{falls } x \in [a, b] \\ -1 & \text{sonst} \end{cases}$$

Entwerfen Sie jeweils einen Lernalgorithmus für das Lernen der Hypothesenklasse \mathcal{H} im nicht-realisierbaren Fall. Ihr Lernalgorithmus sollte den Trainingsfehler minimieren und möglichst effizient sein. Nehmen Sie an, der Lernalgorithmus bekommt als Eingabe ein Sample $S = \{(x_1, y_1), \dots, (x_m, y_m)\}$. Analysieren Sie die Laufzeit in Abhängigkeit von m .