Thomas Kesselheim
Alexander Braun

14. April 2022
Abgabe: 20. April 2022, 9:00 Uhr

Algorithmische Grundlagen des Maschinellen Lernens

Sommersemester 2022

Übungsblatt 2

Aufgabe 1: (6 Punkte)

(Vereinigung von k Intervallen) Betrachten Sie als Hypothesenklasse $\mathcal{H}$ die Menge aller Funktionen der Form

$$h_{a,b} \colon \mathbb{R} \to \{-1, +1\}, \text{ mit } a, b \in \mathbb{R}^k \text{ und}$$

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

Sei $S = \{x_1, \dots, x_m\}$ mit $x_i < x_j$ für $1 \leq i < j \leq m$. Wieviele verschiedene Wege gibt es, den Elementen in S Labels zu geben mithilfe einer Funktion in $\mathcal{H}$? In anderen Worten, wie groß ist $\left|\mathcal{H}|_S\right|$?

Tipp: Betrachten Sie dazu das Ziehen der Intervallgrenzen a_j, b_j für $1 \leq j \leq k$ aus den Intervallen $I_0 = (-\infty, x_1), I_i = (x_i, x_{i+1})$ für $1 \leq i < m$ und $I_m = (x_m, +\infty)$.

Aufgabe 2: (1+2 Punkte)

Sei $\mathcal{R}$ ein Mengensystem mit Grundmenge $\mathcal{X}$ und seien $A, B \subseteq \mathcal{X}$ beliebige Teilmengen. Zeigen Sie, dass

(a) $\mathcal{R}|_{A \cup B}|_A = \mathcal{R}|_A$

(b) $\max\{\dim(\mathcal{R}|_A), \dim(\mathcal{R}|_B)\} \leq \dim(\mathcal{R}|_{A \cup B})$.

Aufgabe 3: (3 Punkte)

Sei $\mathcal{X} = \{1, \dots, n\}$ für ein $n \in \mathbb{N}_{\geq 1}$. Betrachten Sie für ein $0 \leq k \leq n$ das Mengensystem $\mathcal{R} = \{r \subseteq \mathcal{X} \mid |r| = k\}$ bestehend aus allen k-elementigen Teilmengen aus $\mathcal{X}$. Zeigen Sie, dass $\dim(\mathcal{R}) = \min\{k, n - k\}$.

Aufgabe 4: (2+2 Punkte)

Sei die Grundmenge $\mathbb{R}^2$. Zeigen Sie:

(a) Das Mengensystem aller achsenparallelen Quadrate hat VC-Dimension höchstens 3.
(b) Das Mengensystem aller Quadrate hat VC-Dimension mindestens 4.

Aufgabe 5: (4 Punkte)

Sei $\mathcal{R}$ ein Mengensystem mit Grundmenge $\mathcal{X}$ und sei $A \subseteq \mathcal{X}$ eine beliebige Teilmenge. Zeigen Sie, dass $\dim(\mathcal{R}|_A) \leq \dim(\mathcal{R})$.