Thomas Kesselheim
Alexander Braun

02. Juni 2022
Abgabe: 22. Juni 2022, 9:00 Uhr

Algorithmische Grundlagen des Maschinellen Lernens

Sommersemester 2022

Übungsblatt 8

Kommende Woche ist vorlesungsfreie Zeit, am 7. Juni und 9. Juni finden demnach keine Vorlesungen statt. Gleiches gilt für die Tutorien am 9. Juni. Zudem ist am 16. Juni Fronleichnam, hier entfallen Vorlesung und Tutorien ebenfalls.
Dieses Übungsblatt kann daher bis Mittwoch, 22. Juni, abgegeben werden und wird in den Tutorien am 23. Juni besprochen. Übungsblatt 9 erscheint ebenfalls am 23. Juni.

Aufgabe 1: (1+2 Punkte)
Gegeben eine Menge an Datenpunkten $X \subset \mathbb{R}^n$, Labels $Y = \{-1, 1\}$ und $z = (\mathbf{x}, y) \in X \times Y$. Sei $M \subset \mathbb{R}^n$ und $\mathcal{H} = \{h_\mathbf{w} \colon X \to Y \mid \mathbf{w} \in M\}$ die Menge der linearen Klassifikatoren.

(a) Zeigen Sie, dass das Maximum zweier ρ-Lipschitz stetigen Funktionen $f_1, f_2 \colon M \to \mathbb{R}$ wieder ρ-Lipschitz ist.

(b) Zeigen Sie, dass der Hinge-Loss $l^{\text{hinge}}(h_\mathbf{w}, z) = \max\{0, 1 - y \langle \mathbf{w}, \mathbf{x} \rangle\}$ eine $\|\mathbf{x}\|$-Lipschitz stetige Funktion ist.

Aufgabe 2: (2+1 Punkte)
Gegeben $M \subset \mathbb{R}^n$ und zwei Funktionen $f_1, f_2 \colon M \to \mathbb{R}$.

(a) Die Funktion f_1 sei σ_1-stark konvex und die Funktion f_2 sei σ_2-stark konvex. Zeigen Sie, dass $f_1 + f_2$ dann eine $(\sigma_1 + \sigma_2)$-stark konvexe Funktion ist.

(b) Die Funktion f_1 sei σ-stark konvex und f_2 sei konvex. Zeigen Sie, dass $f_1 + f_2$ auch σ-stark konvex ist.

Aufgabe 3: (5 Punkte)
Zeigen Sie, dass für alle Schritte t in AdaBoost gilt

$$\sum_{i : h_t(x_i) \neq y_i} p_i^{(t+1)} = \frac{1}{2} \ .$$

Diese Gleichung gibt auch eine schöne intuitive Erklärung für die Wahl von $p^{(t+1)}$: Es handelt sich um eine Wahrscheinlichkeitsverteilung, sodass h_t nicht besser ist als zufälliges Raten.

Aufgabe 4: (4 Punkte)
Sei $t \in \mathbb{N}$ und sei $d = 2^t$. Zeigen Sie, dass die VC-Dimension der Klasse der Schwellenwertfunktionen in $\mathbb{R}^d$ mindestens $\log_2 d$ ist. Konstruieren Sie dazu eine t-elementige Menge $A \subseteq \mathbb{R}^d$ die durch die Klasse der Schwellenwertfunktionen aufgespalten wird.
Tipp: Wählen Sie eine Menge $A \subseteq \{+1, -1\}^d$.

Aufgabe 5: (5 Punkte)

Seien $\mathcal{H}_1$ und $\mathcal{H}_2$ zwei Hypothesenklassen mit Grundmenge X und mit VC-Dimension d_1 und d_2, mit $3 \leq \mathrm{d}_1, \mathrm{d}_2 < \infty$. Sei $\mathcal{H}'$ die Hypothesenklasse aller Funktionen der Form $g : X \to \{+1, -1\}$ definiert durch $h_1 \in \mathcal{H}_1, h_2 \in \mathcal{H}_2$ mit

$$g(x) = \begin{cases} +1 & \text{falls } h_1(x) = 1 \text{ und } h_2(x) = -1 \\ -1 & \text{sonst} \end{cases}$$

Zeigen Sie eine obere Schranke für die VC-Dimension von $\mathcal{H}'$ in Abhängigkeit von d_1 und d_2.
Tipp: Modifizieren Sie den Beweis von Satz 17.7.