Thomas Kesselheim
Alexander Braun

07. Juli 2022
Abgabe: 13. Juli 2022, 9:00 Uhr

Algorithmische Grundlagen des Maschinellen Lernens

Sommersemester 2022

Übungsblatt 11

Aufgabe 1: (5 Punkte)

Sei $X = \{0,1\}^d$. Wir betrachten die Klasse der Funktionen $\mathcal{F}$ der Form $f_i\colon X \to \{0,1\}$ mit $f_i(x) = x_i$ für $x \in X$ und als Wahrscheinlichkeisverteilung die Gleichverteilung auf $\mathcal{F}$.
Sei $d\colon X \times X \to \mathbb{N}$ definiert durch $d(x,y) = \sum_{i=1}^{d} |x_i - y_i|$.

(a) Zeigen Sie, dass d eine Metrik ist.

(b) Zeigen Sie, dass $\mathcal{F}$ $(\frac{d}{2}, \frac{d}{2}, \frac{1}{2}, \frac{1}{2})$-lokalitätssensitiv ist.

Aufgabe 2: (4+2 Punkte)

Sei wieder $\mathbb{S}^1 = \{\mathbf{x} \in \mathbb{R}^2 \mid \|\mathbf{x}\| = 1\}$. Untersuchen Sie folgenden Abstandsfunktionen $d : \mathbb{S}^1 \times \mathbb{S}^1 \to \mathbb{R}_{\geq 0}$ bezüglich ihrer metrischen Eigenschaften. Zeigen Sie entweder, dass die Abstandsfunktion eine Metrik ist, oder geben Sie ein Beispiel das zeigt, dass mindestens eine der Eigenschaften einer Metrik verletzt wird.

(a) $d(\mathbf{x}, \mathbf{y}) = \cos^{-1} \langle \mathbf{x}, \mathbf{y} \rangle$

(b) $d(\mathbf{x}, \mathbf{y}) = 1 - \langle \mathbf{x}, \mathbf{y} \rangle$

Aufgabe 3: (4 Punkte)

Sei $\alpha = \frac{1+\sqrt{5}}{2}$. Wir untersuchen hierarchisches Clustering bezüglich der k-Center Zielfunktion. Betrachten Sie dazu die Punkte $a = -(2\alpha+1), b = -1, c = 1, d = 2\alpha + 1$. Zeigen Sie, dass es für jedes hierarchische Clustering $\mathcal{C} = \{C_1, C_2, C_3, C_4\}$ auf $S = \{a, b, c, d\}$ ein $i \in \{1,2,3,4\}$ existiert mit

$$\phi(C_i) \geq \alpha \cdot \phi(C_i^*) \ ,$$

wobei C_i^* ein optimales i-Center-Clustering der Menge S ist.