Thomas Kesselheim
Alexander Braun

19. Mai 2022
Abgabe: 01. Juni 2022, 9:00 Uhr

Algorithmische Grundlagen des Maschinellen Lernens

Sommersemester 2022

Übungsblatt 7

Aufgrund von Christi Himmelfahrt entfallen am kommenden Donnerstag, 26. Mai, die Vorlesung und die Tutorien. Dieses Übungsblatt kann daher bis Mittwoch, 1. Juni, abgegeben werden und wird in den Tutorien am 2. Juni besprochen. Übungsblatt 8 erscheint ebenfalls am 2. Juni.

Aufgabe 1: (3+2 Punkte)

Sei $X = \mathbb{R}$, $F = \mathbb{R}^3$ und $\psi\colon X \to F$ mit $\psi(x) = (1, x, x^2)$. Betrachten Sie die Hypothesenklassen $\mathcal{H}_1 = \{h_{a,b,c} \mid a, b \in \mathbb{R}, c \in \{-1, +1\}\}$, $\mathcal{H}_2 = \{h_{\mathbf{w}} \mid \mathbf{w} \in \mathbb{R}^3\}$, wobei

$$h_{a,b,c}(x) = \begin{cases} c & \text{falls } x \in [a, b] \\ -c & \text{sonst} \end{cases} \qquad \text{und} \qquad h_{\mathbf{w}} = \begin{cases} +1 & \text{falls } \langle \mathbf{w}, \psi(x) \rangle \geq 0 \\ -1 & \text{sonst} \end{cases} .$$

(a) Zeigen Sie, dass $\mathcal{H}_1 \subseteq \mathcal{H}_2$.

(b) Zeigen Sie, dass $\mathcal{H}_1 \not\supseteq \mathcal{H}_2$.

Aufgabe 2: (3+2 Punkte)

Sei $X = \mathbb{R}^2$, $F = \mathbb{R}^3$ und $\psi\colon X \to F$ mit $\psi(x_1, x_2) = (x_1, x_2, x_1^2 + x_2^2)$. Betrachten Sie die Hypothesenklassen $\mathcal{H}_1 = \{h_{\mathbf{p},r} \mid \mathbf{p} \in \mathbb{R}^2, r \in \mathbb{R}, r \geq 0\}$ und $\mathcal{H}_2 = \{h_{\mathbf{w},u} \mid \mathbf{w} \in \mathbb{R}^3, u \in \mathbb{R}\}$ wobei

$$h_{\mathbf{p},r}(\mathbf{x}) = \begin{cases} +1 & \text{falls } \|\mathbf{x} - \mathbf{p}\| \leq r \\ -1 & \text{sonst} \end{cases} \qquad h_{\mathbf{w},u}(\mathbf{x}) = \begin{cases} +1 & \text{falls } \langle \mathbf{w}, \psi(\mathbf{x}) \rangle \geq u \\ -1 & \text{sonst} \end{cases}$$

(a) Zeigen Sie, dass $\mathcal{H}_1 \subseteq \mathcal{H}_2$.

(b) Zeigen Sie, dass $\mathcal{H}_1 \not\supseteq \mathcal{H}_2$.

Aufgabe 3: (3+2 Punkte)

Seien $\psi\colon X \to F$ eine Einbettung und $f(\mathbf{w}) = \lambda \|\mathbf{w}\|^2 + \frac{1}{m} \sum_{i=1}^m \max\{0, 1 - y_i \langle \mathbf{w}, \psi(\mathbf{x}_i) \rangle\}$ die Soft-SVM-Zielfunktion auf $(\psi(\mathbf{x}_1), y_1), \ldots, (\psi(\mathbf{x}_m), y_m)$. Betrachten wir nun $\hat{f}\colon \mathbb{R}^m \to \mathbb{R}$ definiert über $\hat{f}(\alpha_1, \ldots, \alpha_m) = f(\sum_{i=1}^m \alpha_i \psi(\mathbf{x}_i))$.

(a) Zeigen Sie, dass $\hat{f}$ konvex ist. Sie dürfen ohne Beweis nutzen, dass f konvex ist.

(b) Drücken Sie $\hat{f}(\alpha_1, \ldots, \alpha_m)$ in Abhängigkeit der zu ψ zugehörigen Kernel-Funktion $K\colon X \times X \to \mathbb{R}$ aus, entfernen Sie dabei alle Vorkommen von $\psi(\mathbf{x}_i)$.

Aufgabe 4: (3+2 Punkte)

Sei S eine Menge von m Datenpunkten mit Labels $(x_1, y_1), \ldots, (x_m, y_m)$, $x_i \in X := \mathbb{R}$, $y_i \in \{-1, +1\}$; $x_i \neq x_j$ für $i \neq j$. Wir betrachten nun die polynomielle Einbettung von $X = \mathbb{R}$ in $F = \mathbb{R}^{k+1}$ mit $\psi(x) = (1, x, x^2, \ldots, x^k)$. Die Menge S wird von einem linearen Klassifikator $h_{\mathbf{w}}$ in F korrekt klassifiziert, wenn $h_{\mathbf{w}}(\psi(x_i)) = y_i$ für alle i, wobei $h_{\mathbf{w}}(\psi(x_i)) = 1$, falls $\langle \mathbf{w}, \psi(x_i) \rangle \geq 0$, -1 sonst.

(a) Zeigen Sie für $m = k + 1$, dass es einen linearen Klassifikator gibt der S korrekt klassifiziert.

(b) Finden Sie für $m = k + 2$ eine Menge S, sodass kein linearer Klassifikator S korrekt klassifziert.

Tipp: (a) Polynominterpolation.