Thomas Kesselheim
Alexander Braun

2. Juni 2022

Algorithmische Grundlagen des Maschinellen Lernens

Sommersemester 2022

Präsenzblatt 8

Aufgabe 1:

Wir betrachten eine Menge von Datenpunkte in $\mathbb{R}$ mit Labels $S = \{(x_1, y_1), \ldots, (x_m, y_m)\} \subseteq \mathbb{R} \times \{-1, 1\}$. Konstruieren Sie eine Hypothese h^* als Linearkombination von Schwellenwertfunktionen, sodass $h^*(x_i) = y_i$ für alle i. Sie dürfen dabei annehmen, dass S keine zwei Punkte enthält mit $x_i = x_j$ und $y_i \neq y_j$. Es sind auch negative Koeffizienten in der Linearkombination erlaubt.

Aufgabe 2:

Sei S eine Menge von m Datenpunkt-/Label-Paaren $(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_m, y_m)$, wobei $\mathbf{x}_i = (x_{i,1}, x_{i,2}) \in \mathbb{R}^2$ und

$$y_i = \begin{cases} +1 & \text{falls } x_{i,1} \geq 0 \text{ und } x_{i,2} \geq 0 \\ -1 & \text{sonst} \end{cases} .$$

Zeigen Sie, dass es für jede solche Menge S einen linearen Klassifikator $h_{\mathbf{w},u}$ ($\mathbf{w} \in \mathbb{R}^2, u \in \mathbb{R}$, $h_{\mathbf{w},u}(x) = 1$, falls $\langle \mathbf{w}, x \rangle \geq u$, -1 sonst) gibt mit $\mathrm{err}_S(h_{\mathbf{w},u}) \leq \frac{1}{3}$. Beachten Sie, dass für $\mathbf{w}$ auch der Nullvektor zugelassen ist.