Anne Driemel
Thomas Kesselheim
Xianghui Zhong

18. Mai 2020
Abgabe bis: 25. Mai 2020, 12 Uhr

Algorithmische Grundlagen des Maschinellen Lernens

Sommersemester 2020

Übungsblatt 5

Aufgabe 1: (3+3 Punkte)
Betrachten Sie den Beweis aus der Vorlesung zur NP-Schwerheit des Lernproblems für die Hypothesenklasse der homogenen Halbräume, $\mathcal{H}_0$.

(a) Geben Sie ein Beispiel für eine Klausel C und eine Hypothese $h_w \in \mathcal{H}_0$, sodass $\phi(C)$ von h_w nicht korrekt klassifiziert wird, aber $\alpha(w)$ die Klausel C erfüllt.

(b) Geben Sie ein Beispiel für eine Klausel C und eine Hypothese $h_w \in \mathcal{H}_0$, sodass $\phi(C)$ von h_w korrekt klassifiziert wird, aber $\alpha(w)$ die Klausel C nicht erfüllt.

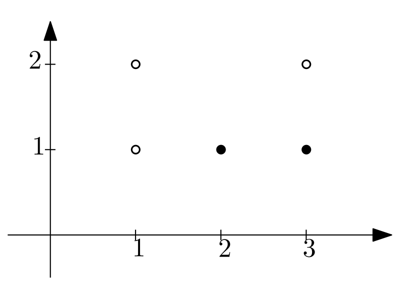
Aufgabe 2: (5 Punkte)
Sei $\mathcal{H}$ die Hypothesenklasse der Halbräume in $\mathbb{R}^2$. Entwerfen Sie einen polynomiellen Lernalgorithmus für $\mathcal{H}$ im nicht-realisierbaren Fall. Beschreiben Sie Ihren Algorithmus und analysieren Sie die Laufzeit. Begründen Sie warum Ihr Algorithmus den Trainingsfehler minimiert.

Aufgabe 3: (2 Punkte)
Sei $\mathcal{H}_z$ für einen Vektor $z \in \mathbb{R}^d$ die Menge von Funktionen der Form $h_{w,u} : \mathbb{R}^d \to \{-1, +1\}$ mit $w \in \mathbb{R}^d, u \in \mathbb{R}$ und

$$h_{w,u}(x) = \begin{cases} +1 \text{ falls } \langle w, x - z \rangle \geq v \\ -1 \text{ sonst} \end{cases}$$

Zeigen Sie, dass diese Hypothesenklasse äquivalent ist zu der in der Vorlesung verwendeten Hypothesenklasse der Halbräume in $\mathbb{R}^d$.

Aufgabe 4: (3 Punkte)
Sei $S = \{((1,1), -1), ((1,2), -1), ((3,2), -1), ((2,1), +1), ((3,1), +1)\}$. Bestimmen Sie manuell eine optimale Hard-SVM-Lösung für S.



Aufgabe 5: (4 Punkte)
Geben Sie ein Beispiel im realisierbaren Fall an, in dem die Soft-SVM-Lösung eine Hypothese liefert, die nicht alle Datenpunkte korrekt klassifiziert.