Hans U. Simon
Christoph Ries

Bochum, den 7. Juni 2016
Abgabe bis zum 14.06.2016

Übungen zur Vorlesung

Theorie des maschinellen Lernens

Sommer 16

Übungsblatt 07

Aufgabe 7.1 (4 Punkte)
Sei $h_w : \mathbb{R}^3 \to [0,1]$ mit $w \in \mathbb{R}^3$. Es gelte

$$h_{(w_1,w_2,w_3)}(x_1,x_2,x_3) := \frac{1}{1+e^{-(w_1x_1+w_2x_2+w_3x_3)}}.$$

Gib für die Trainingsmenge

$$S = \{((-1,2,1),-1),((2,6,1),1),((-2,5,1),1),\\ ((3,0,1),1),((1,0,1),-1),((-3,-1,1),-1)\}$$

ein w an, sodass alle Punkte aus S korrekt klassifiziert werden. Bestimme anschließend die Wahrscheinlichkeiten, die durch deine Wahl von w jedem Punkt aus S zugeordnet werden (vgl. Lemma 9.19 im Skriptteil „Lineare Voraussagefunktionen“ unter Materialien auf der Homepage).

Aufgabe 7.2 (4 Punkte)
$\mathcal{G}_T$ bezeichne die Klasse der Abbildungen von $\mathbb{R}$ nach $\{-1,1\}$, die auf der reellen Zahlengeraden höchstens T-mal zwischen den Werten -1 und 1 alternieren. DS_1 repräsentiere die Klasse der Entscheidungsstümpfe über $\mathbb{R}$ und $LIN(\mathcal{B},T)$ sei die Klasse aller Linearkombinationen von T aus der Basisklasse $\mathcal{B}$ stammenden Hypothesen.
Zeige, dass $\mathcal{G}_T$ eine Teilklasse von $LIN(DS_1,T)$ ist.

Aufgabe 7.3 (4 Punkte)
Es sei die Trainingsmenge

$$S = \{((1,2),2),((4,3),-2),((-2,-1),-2),((5,-2),3),\\ ((6,1),-1),((2,-3),-1),((-1,-4),1)\}$$

gegeben.
Bestimme gemäß der ERM-Regel eine Hypothese aus der Klasse DS_2. Nutze den Algorithmus A_{DS} wie bei Theorem 10.5 im Skriptteil „Boosting“ (siehe unter Materialien auf der Homepage).

— Bitte wenden! —

Aufgabe 7.4 (4 Punkte)

Gegeben sei eine Trainingsmenge $S = \{(x_i, y_i)\}_{i=1}^m \in (\mathcal{X} \times \{-1, 1\})^m$, Hypothesen $g_1, ..., g_T$: $\mathcal{X} \to \{-1, 1\}$, sowie $a_1, ..., a_T \in [0, 1]$ mit $\sum_{i=1}^T a_i = 1$. Es gelte für ein $\theta > 0$, dass

$$y_i \sum_{j=1}^{T} a_j g_j(x_i) \geq \theta$$

für alle $1 \leq i \leq m$.

a) Zeige, dass sich die y_i für alle $1 \leq i \leq m$ durch ein gewichtetes Majoritätskriterium der $g_1, ..., g_T$ bestimmen lassen.

b) Sei D eine Verteilung über $[m]$. Zeige, dass mit obigen Voraussetzungen und $\theta = 2\gamma$ eine Hypothese g_j existiert mit

$$Pr_{i \sim D}[y_i \neq g_j(x_i)] \leq \frac{1}{2} - \gamma.$$