Hans U. Simon
Malte Darnstädt

Bochum, den 04.06.2013
Abgabe bis zum 12.06.2013

Übungen zur Vorlesung

Theorie des Maschinellen Lernens

Sommer 2013

Übungsblatt 07

Aufgabe 7.1

Beweise oder widerlege:

Für alle Klassen F, alle Stichproben z und alle $\gamma > 0$ gilt:
Es gibt eine Funktion $f \in F$, die gleichzeitig $\hat{\mathrm{er}}_z(f)$ und $\hat{\mathrm{er}}_z^\gamma(f)$ minimiert.

Falls die Aussage stimmt: Gib einen Algorithmus an, der f findet.
Falls die Aussage falsch ist: Ist sie dann für alle Klassen F falsch (mit $|F| > 1$)?

Aufgabe 7.2

Beweise folgende obere Schranke:

Sei L ein Algorithmus, der eine Hypothese aus F mit einem kleinsten empirischen γ-Margin-Fehler ausgibt. Dann gilt für alle $m \geq d$, $1 > \delta > 0$ und $\gamma > 0$

$$\epsilon_L(m, \delta, \gamma) \leq \left(\frac{72}{m}\left(d \ln\left(\frac{8em}{d}\right) + \ln\left(\frac{3}{\delta}\right)\right)\right)^{1/2}$$

mit $d = \mathrm{Pdim}(F)$.

Hinweis: Orientiere dich am Beweis von der oberen Schranke, die die Fat-Shattering-Dimension verwendet. Für eine Schranke mit anderen Konstanten gibt es keinen Punktabzug.

Aufgabe 7.3

Zeige, dass der Beweis für die obere Schranke der Covering-Number mit Hilfe der Pseudo-Dimension nicht direkt auf die Graph-Dimension (siehe Aufgabe 6.4) übertragbar ist.

Zeige konkret: Für alle $0 < \epsilon < 1$ existiert eine Klasse F, so dass gilt

$$\infty = \mathrm{Gdim}(Q_\epsilon(F)) > \mathrm{Gdim}(F) = 1$$

Aufgabe 7.4 *(4 Punkte + 4 Bonuspunkte)*

Sei F die Klasse der Funktionen von $[0, 1]$ nach $[0, 1]$ mit einer beschränkten totalen Variation von höchstens V. Sei I die Klasse der monoton steigenden Funktionen von $[0, 1]$ nach $[0, 1 + V/2]$.

Zeige:

a) Für jedes $f \in F$ existieren $g, h \in I$ mit $f = g - h$

b) Für alle $x \in [0, 1]^m$ gilt $\mathcal{N}(\epsilon, F_{|x}, d_1) \leq \mathcal{N}(\frac{\epsilon}{2}, I_{|x}, d_1)^2$

c) *(+4 Bonuspunkte)* $\log_2 \mathcal{N}_1(\epsilon, F, m) = O(V/\epsilon)$