Hans U. Simon
Malte Darnstädt

Bochum, den 09.07.2013
Abgabe bis zum 17.07.2013

Übungen zur Vorlesung

Theorie des Maschinellen Lernens

Sommer 2013

Bonusblatt

Aufgabe 12.1 *(4 Bonuspunkte)*

Folgendes ist aus dem Beweis der oberen Schranke für reelle Klassen bekannt:
Falls für die reelle Funktionenklasse F mit Pseudo-Dimension d ein approximativer SEM-Algorithmus zur Verfügung steht, gibt es einen Lerner L für den mit einer Wahrscheinlichkeit von mindestens $1-\delta$ gilt

$$\begin{aligned} \mathrm{er}_P(L(z)) &\leq \hat{\mathrm{er}}_z(L(z)) + 1/2 \cdot \epsilon_0(m, \delta \,|\, d) \\ &\leq \mathrm{opt}_P(F) + \epsilon_0(m, \delta \,|\, d) \end{aligned}$$

wobei $\epsilon_0(m, \delta \,|\, d) = O\left(\sqrt{\frac{1}{m}(d \ln m + \ln \frac{1}{\delta})}\right)$.

Formuliere und beweise auf der Basis dieser Erkenntnis einen (nicht-trivialen) Satz über strukturelle Risikominimierung im Regressionsmodell:

Sei $F_1 \subsetneq F_2 \subsetneq F_3 \subsetneq \ldots$ eine Folge von reellen Funktionsklassen mit...

Aufgabe 12.2 *(4 Bonuspunkte)*

Sei F eine Klasse von beschränkten Funktionen, die jedoch keine universelle obere Schranke besitzt. Das heißt $\forall f \in F : \exists c : \sup_x |f(x)| \leq c$, aber $\nexists c : \forall f \in F : \sup_x |f(x)| \leq c$.

Zeige, das ein solches F mit endlicher Pseudo-Dimension existiert, dass nicht mit einer polynomiellen Anzahl an Beispielen im Regressionsmodell lernbar ist (polynomiell in $1/\epsilon$, $\log 1/\delta$ und $\mathrm{Pdim}(F)$), selbst wenn die Verteilungen P auf $X \times [0,1]$ beschränkt sind.

Aufgabe 12.3 *(4 Bonuspunkte)*

Sei F eine Menge von Abbildungen von X nach $[0,1]$ und sei P sei eine Verteilung auf $X \times [1-B, B]$ mit $B \geq 1$. Sei $f_a \in \overline{F}$ und $f \in F$.

a) Zeige, dass fast sicher $|y - f_a(x)| \leq B$ gilt.

b) Zeige, dass gilt:

$$E\Big[\Big((y-f(x))^2 - (y-f_a(x))^2\Big)^2\Big] \leq 4B^2\, E\Big[\Big(f_a(x) - f(x)\Big)^2\Big]$$

Lernen mit Äquivalenzanfragen
Sei H_n eine binäre Funktionenklasse über einer Menge X_n und sei $t_n \in H_n$ das unbekannte Zielkonzept.
Ein *EQ-Orakel* arbeitet folgendermaßen: Das Orakel erhält als Eingabe eine Hypothese $h \in H_n$. Gilt $h = t_n$ gibt das Orakel "Ja" zurück; ansonsten antwortet es mit einem Gegenbeispiel $(x, t_n(x))$ mit $h(x) \neq t_n(x)$.
Ein *exakter Lerner mit EQ-Orakel* ist ein Algorithmus der Zugriff auf ein EQ-Orakel hat und als Eingabe n erhält (eine Stichprobe steht nicht zur Verfügung). Der Lerner darf beliebige Anfragen aus H_n an das Orakel stellen. Der Algorithmus terminiert erst dann, wenn das Orakel "Ja" zurückgegeben hat, d.h. am Ende ist das Zielkonzept t_n exakt bekannt.
Ein exakter Lerner mit EQ-Orakel heißt *effizient*, wenn der gesamte Lernvorgang höchstens polynomiel viele Schritte in n benötigt. Dabei gehen wir davon aus, dass das Orakel für seine Antwort nur einen Zeitschritt benötigt.

Aufgabe 12.4 *(4 Bonuspunkte)*

Sei $k \in \mathbb{N}$ fest. Die Klasse k-CNF besteht aus einer Konjunktion von Disjunktionen über booleschen Variablen mit höchstens k Literalen pro Disjunktion. Die Anzahl an Variablen ist der Komplexitätsparameter n. (Bsp für eine 3-CNF über 5 Variabeln: $(x_2 \vee \overline{x}_3 \vee x_5) \wedge x_1 \wedge (\overline{x}_3 \vee \overline{x}_5) \wedge (x_2 \vee x_4)$).

Gib einen effizienten, exakten Lerner mit EQ-Orakel für die Klasse der k-CNF an.

Aufgabe 12.5 *(4 Bonuspunkte)*

Wir nehmen, an dass man bei gegebenem h und x den Wert $h(x)$ effizient berechnen kann.

Zeige: Wenn für $\cup_n H_n$ ein effizienter, exakter Lerner mit EQ-Orakel existiert, dann existiert ein effizienter Lerner für $\cup_n H_n$ im binären, eingeschränkten Model.