

Übungen zur Vorlesung
Theorie des Maschinellen Lernens
Sommer 2014
Übungsblatt 06

Aufgabe 6.1

Seien $\mathcal{H}_1, \dots, \mathcal{H}_s$ Basisklassen mit VC-Dimensionen d_1, \dots, d_s . Sei NN_s die Klasse der Neuronalen Netzwerke mit s Berechnungsknoten, wobei die Funktion von Knoten i aus Basisklasse \mathcal{H}_i stammt.

Beweise die Schranke der Kapazitätsfunktion für NN_s . Zeige also:

$$\Pi_{\text{NN}_s}(m) \leq \left(\frac{em}{d_1}\right)^{d_1} \cdots \left(\frac{em}{d_s}\right)^{d_s}$$

Aufgabe 6.2

Sei \mathcal{C} eine Konzeptklasse über X mit VC-Dimension d . Ein Algorithmus heißt *schwacher PAC-Lerner für \mathcal{C}* wenn ein festes $\gamma > 0$ existiert, so dass für alle Zielverteilungen D über X , alle Zielkonzepte $c \in \mathcal{C}$ und alle $\delta > 0$ der Algorithmus mit Zugriff auf ein Beispielorakel bei Eingabe von δ eine Hypothese ausgibt, die mit Wahrscheinlichkeit $\geq 1 - \delta$ einen Fehler kleiner als $1/2 - \gamma$ besitzt.

Zeige: Jeder schwache Lerner für \mathcal{C} benötigt mindestens eine Stichprobe der Größe

$$m = \Omega(d\gamma).$$

Hinweis: Da die zu beweisende Schranke von δ unabhängig ist, kann man δ auf einen festen Wert setzen. Es gibt Teilpunkte, wenn man zumindest begründen kann, warum für Klassen mit unendlicher VC-Dimension keine schwachen PAC-Lerner existieren, die mit einer endlichen Stichprobe auskommen.

Aufgabe 6.3

Sei $N = 145$. Die Umkehrabbildung von $f(x) = x^3 \pmod N$ lässt sich in der Form $f^{-1}(y) = y^k \pmod N$ darstellen. Bestimme den Exponenten k und berechne dann $f^{-1}(4)$ mit Hilfe des iterierten Quadrierens.

Aufgabe 6.4

Sei \mathcal{C} eine Konzeptklasse über X .

Zeige: Wenn ein effizienter Vorhersage-Algorithmus für \mathcal{C} existiert, dann existiert auch ein effizienter PAC-Lerner für \mathcal{C} .

Hinweis: Sei A der Vorhersage-Algorithmus. A bildet also gelabelte Stichproben zusammen mit einem ungelabelten Testpunkt auf eine Vorsage des Labels des Testpunktes ab:

$$A : (X \times \{0, 1\})^* \times X \longrightarrow \{0, 1\}$$

Für eine feste Stichprobe T können wir $A_T(x) := A(T, x)$ als eine Abbildung von X nach $\{0, 1\}$ betrachten. Dann wähle als Hypothesenklasse \mathcal{H} des PAC-Lerners:

$$\mathcal{H} = \{A_T \mid T \text{ ist gelabelte Stichprobe}\}$$