

Übungen zur Vorlesung
Theorie des Maschinellen Lernens
Sommer 2014
Übungsblatt 12

Aufgabe 12.1 (4 Bonuspunkte)

Zeige: Wenn eine Klasse \mathcal{T} mit $\mathcal{H} = \mathcal{T}$ effizient agnostisch PAC-lernbar ist, dann ist sie auch effizient unter Klassifikationsrauschen PAC-lernbar.

Aufgabe 12.2 (4 Bonuspunkte)

Sei $X = \mathbb{R}^2$ und $\mathcal{T} = \mathcal{H}$ die Klasse der achsen-parallelen Rechtecke. Wir zeigen, dass diese Klasse agnostisch effizient PAC-lernbar ist.

Gib dazu eine effiziente Methode an, die zu einer Menge $(x_1, b_1), \dots, (x_m, b_m)$ mit $x_i \in \mathbb{R}^2$, $b_i \in \{0, 1\}$ ein Rechteck mit der größten Übereinstimmung findet.

Aufgabe 12.3 (4 Bonuspunkte)

Sei \mathcal{D} eine Verteilung über X . Sei \mathcal{H} eine Klasse probabilistischer Konzepte über X , und $f : X \rightarrow [0, 1]$ ein probabilistisches Zielkonzept.

Zeige: Eine Hypothese $h \in \mathcal{H}$, die die quadratische Distanz

$$E_{x \sim \mathcal{D}}[(f(x) - h(x))^2]$$

minimiert, minimiert ebenfalls den erwarteten quadratischen Verlust

$$E_{(x,y) \sim A_{\mathcal{D},f}}[(h(x) - y)^2].$$

Hinweis: Nimm der Einfachheit halber an, dass X endlich ist.

Aufgabe 12.4 (4 Bonuspunkte)

Sei $\mathcal{H}_n = \mathcal{T}_n$ die Klasse der Perzeptrons mit binären Gewichten über $X_n = \mathbb{R}^n$, die folgendermaßen definiert ist:

Jeder Vektor $u \in \{0, 1\}^n$ und jedes $t \in \mathbb{R}$ definiert ein Konzept $c_{u,t}$ mit

$$c_{u,t}(x) = \begin{cases} 1, & \text{falls } \sum_{i=1}^n u_i \cdot x_i \geq t \\ 0, & \text{sonst} \end{cases}$$

Zeige, dass $\mathcal{H} = (\mathcal{H}_n)_{n \geq 1}$ mit $\mathcal{T} = \mathcal{H}$ nicht effizient im agnostischen PAC-Modell lernbar ist, falls $RP \neq NP$.

Hinweis: Aus der Vorlesung ist bekannt, dass die Klasse der Monome nicht effizient agnostisch PAC-Lernbar ist, falls $RP \neq NP$. Gib eine Reduktion von dem Minimum-Disagreement-Problem für Monome auf das Minimum-Disagreement-Problem für die Perzeptrons an. Versuche dies zuerst für ein festes t (z.B. $t = 1/2$) und überlege dann wie es mit allgemeinem t funktioniert.