

Übungen zur Vorlesung
Theorie des maschinellen Lernens
Sommer 17
Übungsblatt 04

Aufgabe 4.1 (4 Punkte)

Seien $\mathcal{H}_1, \dots, \mathcal{H}_r$ Hypothesenklassen über dem Grundraum X . Es sei $d := \max_i \text{VCdim}(\mathcal{H}_i)$. Der Einfachheit halber setzen wir voraus, dass $d \geq 3$. Zeige, dass

$$\text{VCdim}\left(\bigcup_{i=1}^r \mathcal{H}_i\right) \leq 4d \log(2d) + 2 \log(r).$$

Hinweis Es kann benutzt werden, dass für $a \geq 1$ und $b > 0$ gilt:

$$x \geq 4a \log(2a) + 2b \quad \Rightarrow \quad x \geq a \log(x) + b$$

Aufgabe 4.2 (4 Punkte)

Sei \mathcal{H} eine endliche Hypothesenklasse. Sei $d : \mathcal{H} \rightarrow \{0, 1\}^*$ eine Codierung der Hypothesen. Zeige, dass die VC-Dimension von \mathcal{H} durch die größte Beschreibungslänge einer Hypothese in \mathcal{H} begrenzt wird, genauer dass gilt

$$\text{VCdim}(\mathcal{H}) \leq 2 \sup\{|d(h)| : h \in \mathcal{H}\}.$$

Zeige weiter, dass

$$\text{VCdim}(\mathcal{H}) \leq \sup\{|d(h)| : h \in \mathcal{H}\},$$

falls d präfixfrei ist.

Aufgabe 4.3 (4 Punkte)

a) Sei $\mathcal{H} = \cup_{n=1}^{\infty} \mathcal{H}_n$ eine Hypothesenklasse, wobei \mathcal{H}_n für jedes $n \in \mathbb{N}$ endlich ist. Finde eine Gewichtsfunktion $w : \mathcal{H} \rightarrow [0, 1]$ für die folgendes gilt:

- $\sum_{h \in \mathcal{H}} w(h) \leq 1$
- $\forall h \in \mathcal{H}$: $w(h)$ hängt von $n(h) = \min\{n \mid h \in \mathcal{H}_n\}$ und $|\mathcal{H}_n|$ ab.

b) Gib nun eine Gewichtsfunktion für den Fall an, dass \mathcal{H}_n für alle $n \in \mathbb{N}$ abzählbar ist.

Aufgabe 4.4 (4 Punkte)

Sei $\mathcal{H} \subset \{0, 1\}^{\mathbb{R}} := \{f \mid f : \mathbb{R} \rightarrow \{0, 1\}\}$ die Hypothesenklasse bestehend aus den Vereinigungen von abzählbar unendlich vielen Intervallen. Zeige, dass \mathcal{H} nichtuniform lernbar ist (ohne Effizienzbetrachtung) und bestimme die von SRM verwendete Kostenfunktion.