Computational Learning Theory

Sommersemester 2018

Mario Holldack, M. Sc.
Prof. Dr. Georg Schnitger
Hannes Seiwert, M. Sc.

GOETHE UNIVERSITÄT FRANKFURT AM MAIN

Institut für Informatik
AG Theoretische Informatik

Übungsblatt 2

Ausgabe: 23.04.2018
Abgabe: 30.04.2018, 12:15

Wählen Sie drei Aufgaben aus, die Sie bearbeiten. Die Maximalpunktzahl dieses Blattes beträgt 32 Punkte.

Aufgabe 2.1 *Agnostisches Lernen: Lineare Regression* (12 Punkte)

Sei $t \in \mathbb{R}_{>0}$, $X = [0, t]$ und $Y = \mathbb{R}$. Wir betrachten die Hypothesenklasse $\mathcal{H}$ aller affinen Funktionen $f : X \to Y$ mit $f(x) = a \cdot x + b$, wobei $a, b \in \mathbb{R}$.

Sei $f \in \mathcal{H}$ mit $f(x) = a \cdot x + b$ im Folgenden das Zielkonzept. Wir wollen f anhand von Beispielen lernen. Die Beispiele sind aber durch einen additiven Fehler ξ verrauscht, wobei ξ normalverteilt um den Mittelwert $\mu = 0$ mit Varianz $\sigma^2 > 0$ sei. Der Lernalgorithmus erhält nur die verrauschten Beispiele (x, y) mit $y = f(x) + \xi$.

Die Verteilung D über $X \times Y$ sieht folgendermaßen aus: $D(x, \cdot)$ ist die Gleichverteilung auf X und für jedes gegebene $x \in X$ ist $D(x, y|x)$ die Normalverteilung um den Mittelwert $\mu = f(x)$ mit Varianz σ^2. Als Loss-Funktion wählen wir den quadratischen Loss.

a) Sei die Beispielmenge $S = \{(x_1, y_1), \ldots, (x_s, y_s)\}$ gegeben und seien $\overline{x} = \frac{1}{s}\sum_{i=1}^{s} x_i$ bzw. $\overline{y} = \frac{1}{s}\sum_{i=1}^{s} y_i$ die (empirischen) Mittelwerte der Beispielmenge.

   Zeigen Sie, dass die Hypothese h_S mit

   $$h_S(x) = \hat{a} \cdot x + \hat{b} \quad \text{mit} \quad \hat{a} = \frac{\mathrm{Cov}[x, y]}{\mathrm{Var}[x]} \quad \text{und} \quad \hat{b} = \overline{y} - \hat{a} \cdot \overline{x}$$

   eine ERM-Hypothese bzgl. des quadratischen Losses ist. Dabei ist $\mathrm{Var}[x] = \frac{1}{s}\sum_i (x_i - \overline{x})^2$ die (empirische) Varianz und $\mathrm{Cov}[x, y] = \frac{1}{s}\sum_i (x_i - \overline{x})(y_i - \overline{y})$ die (empirische) Kovarianz der Beispielmenge.

   Hinweis: $\mathrm{Loss}^S(h_S)$ ist eine reellwertige Funktion, die von zwei reellen Parametern abhängt, also $\mathrm{Loss}^S(h_S) : \mathbb{R}^2 \to \mathbb{R}$. Wie minimiert man eine solche Funktion?

b) Berechnen Sie den Approximationsfehler von $\mathcal{H}$. Berechnen Sie anschließend den Schätzfehler einer gegebenen Hypothese $h \in \mathcal{H}$ mit $h(x) = a' \cdot x + b'$ in Abhängigkeit von $(a' - a)$, $(b' - b)$ und t.

   Zur Erinnerung: Das Zielkonzept ist $f \in \mathcal{H}$ mit $f(x) = a \cdot x + b$.

Aufgabe 2.2 *VC-Dimension von konvexen Polygonen* (10 Punkte)

Sei $k \in \mathbb{N}_{\geq 3}$ und $X = \mathbb{R}^2$ der Beispielraum. Die Konzeptklasse k-GONS bestehe aus allen konvexen Polygonen in $\mathbb{R}^2$, die höchstens k Ecken besitzen. Das Konzept eines Polygons sei die konvexe Hülle seiner Ecken. Ein solches Polygon kann z.B. dargestellt werden als der Durchschnitt von k Halbebenen.

Zeigen Sie $\mathrm{VC}(k\text{-GONS}) = 2k + 1$

Bitte wenden!

Aufgabe 2.3 *VC-Dimension von Rechtecken* (10 Punkte)

Sei $d \in \mathbb{N}_{>0}$ und $X = \mathbb{R}^d$ der Beispielraum. Für Vektoren $a{=}(a_1, \dots, a_d), b{=}(b_1, \dots, b_d) \in \mathbb{R}^d$ schreiben wir $a < b$ (bzw. $a \leq b$), falls $a_i < b_i$ (bzw. $a_i \leq b_i$) für *alle* $i = 1, \dots, d$ gilt.

Für zwei Vektoren $a, b \in \mathbb{R}^d$ mit $a < b$ sei $R_{a,b} := \{x \in \mathbb{R}^d : a \leq x \leq b\}$ das von ihnen aufgespannte (achsenparallele) Rechteck mit den Eckpunkten a, b.

a) Die Konzeptklasse $\text{RECHTECK}_d = \{R_{a,b} : a, b \in \mathbb{R}^d, a < b\}$ bestehe aus allen d-dimensionalen Rechtecken.

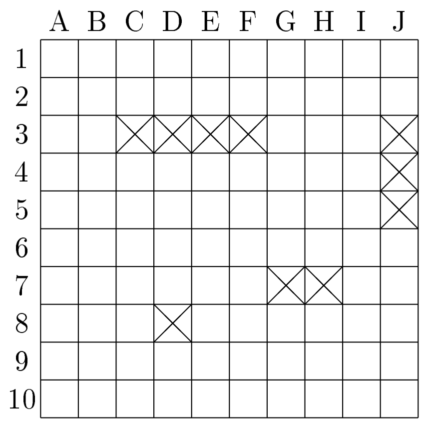
   Zeigen Sie $\text{VC}(\text{RECHTECK}_d) = 2d$.

b) Ein Rechteck $R_{a,b}$ ist ein *Quadrat*, falls alle Kantenlängen gleich sind, d.h. wenn es ein $k \in \mathbb{R}_{>0}$ gibt, sodass $b_i - a_i = k$ für alle i.
   Die Konzeptklasse $\text{QUADRAT}_d = \{R_{a,b} : a, b \in \mathbb{R}^d,\ R_{a,b} \text{ ist ein Quadrat}\}$ bestehe aus allen d-dimensionalen Quadraten.

   Bestimmen Sie die VC-Dimension von QUADRAT_d.

Aufgabe 2.4 *Schiffe versenken* (10 Punkte)

Wir betrachten das Spiel *Schiffe versenken*. Gespielt wird auf einem Spielfeld $F = \{1, \dots, 10\} \times \{A, \dots, J\}$ mit 10 mal 10 Zellen. Ein Spieler hat vier Schiffe mit den jeweiligen Größen $1{\times}4$, $1{\times}3$, $1{\times}2$ und $1{\times}1$. Er darf die Schiffe auf den Zellen beliebig waagerecht oder senkrecht (nicht diagonal) positionieren, wobei sich die Schiffe in keiner Weise berühren dürfen (auch nicht diagonal).



Beispiel: das Konzept $c = \{C3, D3, E3, F3, J3, J4, J5, G7, H7, D8\}$

Der Beispielraum X bestehe aus allen Zellen des Spielfeldes $F = \{1, \dots, 10\} \times \{A, \dots, J\}$. Für eine legale Anordnung der vier Schiffe bestehe deren Konzept aus der Menge der von den Schiffen abgedeckten Zellen. Die Konzeptklasse SCHIFFE-VERSENKEN bestehe aus allen Konzepten von legalen Anordnungen der vier Schiffe.

Geben Sie eine möglichst gute untere Schranke sowie eine möglichst gute obere Schranke für die VC-Dimension dieser Konzeptklasse an.