Anne Driemel
Thomas Kesselheim
Xianghui Zhong

21. April 2020
Abgabe bis: 27. April 2020, 12 Uhr

Algorithmische Grundlagen des Maschinellen Lernens

Sommersemester 2020

Übungsblatt 1

Aufgabe 1: (4 Punkte)
Betrachten Sie einen zufälligen Würfelwurf mit 3 Würfeln. Nehmen Sie an, dass jeder einzelne Würfel gleichverteilt eine Zahl zwischen 1 und 6 liefert und dass die Würfel statistisch unabhängig voneinander sind. Was ist die Wahrscheinlichkeit, dass der Wurf genau zwei gleiche Zahlen enthält?

Aufgabe 2: (4 Punkte)
Zwei Punktmengen im $\mathbb{R}^2$ sind durch eine Gerade separierbar, wenn alle Punkte der einen Menge unterhalb der Gerade liegen und alle Punkte der anderen Menge oberhalb der Geraden liegen. Gleiches gilt in Bezug auf Ebenen im $\mathbb{R}^3$.

1. Geben Sie ein Beispiel zweier disjunkter Punktmengen $P \subset \mathbb{R}^2$ und $Q \subset \mathbb{R}^2$, die nicht durch eine Gerade separierbar sind.

2. Geben Sie ein Beispiel zweier disjunkter Punktmengen $P \subset \mathbb{R}^3$ und $Q \subset \mathbb{R}^3$, die nicht durch eine Ebene separierbar sind.

Begründen Sie Ihre Antworten. Kann in einem Ihrer Beispiele ein Punkt entfernt werden, ohne dass das Beispiel separierbar wird?

Aufgabe 3: (10 Punkte)
Betrachten Sie den Hypothesenraum $\mathcal{H}$, in der jede Funktion $h_{a,b} \colon \mathbb{R} \to \{-1, +1\}$ durch ein *offenes* Intervall (a, b) definiert ist und $h_{a,b}(x) = 1$ genau dann wenn $x \in (a, b)$. Für den Fall, dass $a = b$, definieren wir, dass $(a, b) = \emptyset$ gilt. Betrachten Sie den folgenden Lernalgorithmus.

1. Falls das Sample S kein Element (x_i, y_i) mit $y_i = 1$ enthält, gibt der Lernalgorithmus die konstante Funktion $x \mapsto -1$ für all $x \in \mathbb{R}$ zurück.

2. Für ein Sample $\{(x_1, y_1), \ldots, (x_m, y_m)\}$ gibt der Lernalgorithmus die Funktion $h_{a',b'}$ zurück, die wie folgt definiert ist. Der Wert a' ist das größte x_i, sodass $y_i = -1$ und $x_i < x_j$ für ein $y_j = 1$. Falls dieses nicht existiert, ist $a' = -\infty$. Der Wert b' ist das kleinste x_i, sodass $y_i = -1$ und $x_i > x_j$ für ein $y_j = 1$. Falls dieses nicht existiert, ist $b' = \infty$.

Beweisen Sie, dass dieser Algorithmus die PAC-Lernbarkeit nachweist. Sie können dazu den Beweis aus der Vorlesung modifizieren und gegebenenfalls erweitern.

Aufgabe 4: (2 Punkte)
Sei für eine Hypothese $h \colon \mathbb{R} \to \{-1, +1\}$, $h(x) = 1$ genau dann wenn $x \in [2, 4]$. Berechnen Sie den tatsächlichen Fehler $\mathrm{err}_{\mathcal{D},f}(h)$ unter der Annahme, dass für Grundwahrheit f gilt, dass $f(x) = 1$ genau dann wenn $x \in [1, 3]$ und, dass die Verteilung $\mathcal{D}$ durch die Dichtefunktion $g(x) = e^{-x}$ gegeben ist.