Thomas Kesselheim
Alexander Braun

07. April 2022

Algorithmische Grundlagen des Maschinellen Lernens

Sommersemester 2022

Präsenzblatt 1

Sobald Sie sich in Gruppen eingefunden haben, können Sie gerne mit einer kurzen Vorstellungsrunde beginnen: Wie heißen Sie? Haben Sie Vorwissen im Bereich des Maschinellen Lernens? Studieren Sie Informatik, Mathematik oder etwas anderes?

Anschließend sollten Sie die Aufgaben auf diesem Präsenzübungsblatt diskutieren. Bitte sehen Sie dies auch als Chance, Definitionen, Theoreme und Konzepte der Vorlesung zu wiederholen und zu besprechen. Im weiteren Semester werden die Präsenzaufgaben daraufhin ausgelegt, Diskussionsgrundlage für eingeführte Konzepte und Methoden zu sein. Der Fokus sollte also insbesondere nicht nur auf der Ausarbeitung einer formal korrekten Lösung liegen. Wenn Sie eine Definition oder ein Theorem nicht mehr wissen, schlagen Sie gerne im Skript oder in der Literatur nach.

Sollte am Ende des Tutoriums noch Zeit sein, werden wir gemeinsam Ihre Ideen und Ansätze mit der gesamten Gruppe besprechen.

Aufgabe 1:
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:
Betrachten Sie nun einen zufälligen Würfelwurf mit 2 Würfeln. Nehmen Sie an, dass jeder einzelne Würfel gleichverteilt eine Zahl zwischen 1 und 6 liefert und dass die Würfel statistisch unabhängig voneinander sind.
Sei X_1 die Zufallsvariable, welche das Ergebnis des ersten Würfels ausgibt, und X_2 die Zufallsvariable, welche das Ergebnis des zweiten Würfels ausgibt. Außerdem sei X_3 die Zufallsvariable, welche die Summe der beiden Ergebnisse ausgibt.
Entscheiden sie für je zwei der drei Zufallsvariablen, ob sie identisch verteilt und/oder unabhängig sind.

Aufgabe 3:
Sei $\lambda > 0$, wir betrachten die folgende Dichtefunktion auf $\mathbb{R}$

$$g(x) = \begin{cases} \lambda e^{-\lambda x} & \text{für } x \geq 0 \\ 0 & \text{sonst.} \end{cases}$$

Die Dichtefunktion g definiert eine Wahrscheinlichkeitsverteilung auf $\mathbb{R}$. Die Wahrscheinlichkeit des Intervalls $[a, b] \in \mathbb{R}$ ist dabei durch

$$P([a, b]) = \int_a^b g(x)dx$$

gegeben. Berechnen Sie die Wahrscheinlichkeit, des Intervalls $[-2, 6]$.

Aufgabe 4:
Betrachten Sie die Grundwahrheit $f \colon \mathbb{R}^2 \to \{+1, -1\}$ gegeben durch

$$f(x_1, x_2) = \begin{cases} -1 & \text{falls } 3x_1 + 1 - x_2 < 0 \\ +1 & \text{falls } 3x_1 + 1 - x_2 \geq 0 \end{cases}$$

Stellen Sie zu dieser Grundwahrheit einen geeigneten Hypothesenraum auf.

Aufgabe 5:
(Erwarteter Trainingsfehler) Gegeben die Menge der zu klassifizierenden Datenpunkte X und eine Wahrscheinlichkeitsverteilung $\mathcal{D}$ auf X. Sei $\mathcal{H}$ eine Hypothesenklasse und $f \in \mathcal{H}$ die Grundwahrheit.
Für die Menge $S = \{(x_1, y_1), \ldots, (x_m, y_m)\}$ von m unabhängig und identisch verteilt aus $\mathcal{D}$ gezogenen Samplen bezeichnet

$$\mathrm{err}_S(h) = \frac{1}{m}|\{h(x_i) \neq y_i\}|$$

den Trainingsfehler der Hypothese h auf S. Zeigen Sie, dass der erwartete Trainingsfehler von h dem tatsächlichen Fehler von h entspricht, also

$$\mathbb{E}_{S \sim \mathcal{D}}[\mathrm{err}_S(h)] = \mathrm{err}_{\mathcal{D},f}(h).$$