

Computational Learning Theory

Wintersemester 2011/2012

Prof. Dr. Georg Schnitger

Arbeitsgruppe Theoretische Informatik, Institut für Informatik



Blatt 11

Ausgabe: 30.01.2012

Abgabe: 06.02.2012

11.1. Aufgabe (8)

Maximum-Likelihood

Eine Zufallsquelle A gibt in jedem Schritt eine von n möglichen Ausgaben a_1, \dots, a_n . Die entsprechenden Wahrscheinlichkeiten p_1, \dots, p_n sind uns unbekannt. Wir beobachten s Ausgaben von A . Gib eine ML-Hypothese, also eine Verteilung auf a_1, \dots, a_n , die die Stichprobe am wahrscheinlichsten generiert hat, an. Begründe Deine Antwort.

Hinweis: Für die Verteilungen $(p_i)_{i \in [n]}$, $(q_i)_{i \in [n]}$ gilt die Kullback-Leibler-Divergenz $\sum_{i=1}^n p_i \ln(\frac{q_i}{p_i}) \leq 0$.

11.2. Aufgabe (8)

Maximum-Likelihood

Für die Fernseherfirma F lief das Weihnachtsgeschäft gut. Doch Mitte Januar bekommt F eine Reihe von Reklamationen r_1, \dots, r_k . Die Kunden geben an, dass ihr Fernseher nach einem Versuch des Programmwechsels kapitulierte. Aus internen Speichern kann F ablesen, dass für das Gerät aus Reklamation r_i insgesamt n_i Programmwechsel erfolgreich waren. F will aus den vorliegenden Daten ermitteln, mit welcher Wahrscheinlichkeit p ein Fernseher beim Programmwechsel zerstört wird.

- Welche Wahrscheinlichkeitsverteilung liegt dem Ereignis zugrunde, dass ein Fernseher zum Zeitpunkt t kapituliert?
- Wie lautet die ML-Schätzung für p ?

11.3. Aufgabe (8)

Größe von Entscheidungsbäumen

Knoten einfacher Entscheidungsbäume verzweigen aufgrund des Werts einer Variablen aus einer Variablenmenge $X = \{x_1, \dots, x_n\}$. Wir vergleichen diese Bäume mit Threshold-Entscheidungsbäumen, die aufgrund von Threshold-Funktionen über X verzweigen.

- In der Vorlesung wurde erwähnt, dass einfache Entscheidungsbäume die Größe $\Omega(2^n)$ haben müssen um $x_1 \oplus x_2 \oplus \dots \oplus x_n$ berechnen zu können. Warum ist das richtig?
- Konstruiere einen Threshold-Entscheidungsbaum mit möglichst wenigen Knoten, der genau die Eingaben ungerader Parität akzeptiert.

Fazit: Die Ausdrucksstärke kleiner Entscheidungsbäume hängt entscheidend von der Wahl der Entscheidungsfunktionen ab. Aber, je komplexer die Entscheidungsfunktionen, umso schwieriger wird das Erlernen der entsprechenden Bäume.