

Grundlagen des maschinellen Lernens

Tobias Lang Mathias Schickel Andreas Schilling Sommersemester 2017

(9 Punkte)

Übungsblatt 6

Ausgabe: 16.06.2017; Abgabe: bis 29.06.2017, 23:59 Uhr

Aufgabe 1 (Fragen zu Bayesschen Netzen und Markov-Modellen) (12 Punkte)

Man beantworte folgende Fragen zu *Bayesschen Netzen*. (3 Punkte)

- a) Was repräsentieren die Knoten und Kanten eines Bayesschen Netzes?
- b) Die Marginalisierung spielt bei Bayesschen Netzen eine wichtige Rolle. Zur Lösung welcher Probleme wird sie verwendet und welche zwei wichtigen Begriffe sind mit ihr assoziiert? (2 Punkte)

Ebenso beantworte man folgende Fragen zu Markov-Modellen.

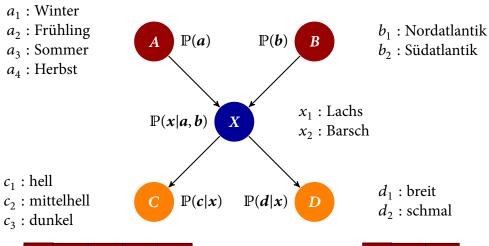
- c) Was versteht man unter der Ordnung eines Hidden Markov Models (HMM)?
- d) Was bedeuten die mit Ω und V bezeichneten Mengen in einem HMM?
- e) Was beschreiben dabei die Matrizen A und B?
- f) Bei HMMs treten drei interessante Probleme auf. Welche sind das und um was geht es bei ihnen jeweils? (3 Punkte)
- g) Wozu werden der Forward- und Backward-Algorithmus verwendet? (2 Punkte)
- h) Warum macht es *keinen* Unterschied, wenn man beim *Viterbi-Dekoder* mit dem Logarithmus rechnet?

Aufgabe 2 (Bayessche Netze – Berechnung)

(10 Punkte)

Man berechne die folgenden Wahrscheinlichkeiten zu dem unten abgebildeten Bayesschen Netz schriftlich (und gebe den Rechenweg an):

- a) $\mathbb{P}(X = \text{Lachs}),$
- b) $\mathbb{P}(A = \text{Winter}, B = \text{Norden}, X = \text{Lachs}, C = \text{hell}, D = \text{schmal}),$
- c) $\mathbb{P}(X = \text{Lachs} | C = \text{mittelhell}),$
- d) $\mathbb{P}(A = \text{Sommer} \mid D = \text{breit}, X = \text{Barsch}),$
- e) $\mathbb{P}(C = \text{hell} \mid D = \text{schmal}, X = \text{Lachs}).$



$\mathbb{D}(a)$	$\mathbb{P}(a_1)$	$\mathbb{P}(a_2)$	$\mathbb{P}(a_3)$	$\mathbb{P}(a_4)$
$\mathbb{P}(a)$	0.25	0.25	0.25	0.25

$\mathbb{P}(m{b})$	$\mathbb{P}(b_1)$	$\mathbb{P}(b_2)$
	0.6	0.4

$\mathbb{P}(x a,b)$	$\mathbb{P}(x_1 a_i,b_j)$	$\mathbb{P}(x_2 a_i,b_j)$
a_1, b_1	0.5	0.5
a_1, b_2	0.7	0.3
a_2, b_1	0.6	0.4
a_2, b_2	0.8	0.2
a_3, b_1	0.4	0.6
a_3, b_2	0.1	0.9
a_4, b_1	0.2	0.8
a_4, b_2	0.3	0.7

	$\mathbb{P}(c x)$	$\mathbb{P}(c_1 x_k)$	$\mathbb{P}(c_2 x_k)$	$\mathbb{P}(c_3 x_k)$
	x_1	0.6	0.2	0.2
I	x_2	0.2	0.3	0.5

$\mathbb{P}(d x)$	$\mathbb{P}(d_1 x_k)$	$\mathbb{P}(d_2 x_k)$
x_1	0.3	0.7
x_2	0.6	0.4

(10 Punkte)

Es sei wieder das Bayessche Netz aus der letzten Aufgabenstellung gegeben. Dieses Netz soll in Matlab mittels einer geeigneten Funktion dargestellt werden. Die Funktion sollte dazu zur Eingabe eines Zustands (A, B, C, D, X) dessen Wahrscheinlichkeit berechnen. Im Falle, dass eine der den Zustand spezifizierenden Variablen nicht gegeben ist, sollen alle möglichen Zustände, die durch die festgelegten Variablen gegeben sind, betrachtet werden.

Hinweise:

- Im Code soll angegeben werden, wie die Zustände dargestellt werden (Zahlen, Strings, etc.).
- Es kann geschickt sein, die Tabellen direkt als Matrizen darzustellen und die Parameter A, B, C, D, X als Zahlen, sodass die Parameter als Indizes verwendet werden können.
- Als Test kann folgender Fall dienen: $\mathbb{P}(Barsch, schmal, Sommer, Nordatlantik) = 0.036$.