

Maschinelles Lernen und Data Mining
 Sommersemester 2010
Übungsblatt 6

Besprechung des Übungsblattes am 24.06.2010

Aufgabe 6-1 Wahrscheinlichkeitsrechnung
schriftlich bearbeiten

Gegeben seien die folgende Daten:

		Y		
		1	2	3
X	1	0,1	0,15	0,25
	2	0,05	0,3	0,15

Berechnen Sie

- die Randwahrscheinlichkeiten $P(X = x_i)$ und $P(Y = y_i)$
- die Erwartungswerte $E(X)$, $E(Y)$
- die Varianzen $Var(X)$, $Var(Y)$, sowie die Kovarianz $Cov(X,Y)$.
- die Korrelation $\rho = \frac{Cov(X,Y)}{\sqrt{Var(X) \cdot Var(Y)}}$ (Achtung, im Skript fehlt die Wurzel des Nenners)
- Sind die Variablen X,Y unabhängig?

Aufgabe 6-2 Bedingte Wahrscheinlichkeiten
schriftlich bearbeiten

Gegeben die Datentabelle aus Aufgabe 1, sowie die folgenden Ereignisse:

$$A \begin{cases} 0 & x \leq y \\ 1 & \text{sonst} \end{cases}$$

$$B \begin{cases} 0 & x = y \\ 1 & \text{sonst} \end{cases}$$

Berechnen Sie: $P(B = 0|A = 0)$, $P(A = 0|B = 0)$, $P(B = 1|A = 0)$

Aufgabe 6-3 Gradientenabstieg
schriftlich

Gegeben die Funktion $f(x) = 2 + 1/8 (x - 5)^2 + \sin(x)$.

Ermitteln sie ein lokales Minimum der Funktion im Bereich $x \in [0, 15]$ mit Hilfe des Gradientenverfahrens. Starten sie bei $x = 14$ und passen Sie die Schrittweite mit dem Faktor $\alpha = \nabla f(x)$; $\alpha = 0,2/1/2$ an.

- (a) Skizzieren sie die Lösung grafisch.
- (b) Beschreiben Sie ihre Beobachtungen.
- (c) Skizzieren Sie mögliche Problemlösungen.

Aufgabe 6-4 Spamfilter

Skizzieren Sie einen einfachen Bayes-Spamfilter. Gehen Sie davon aus, dass die Wörter der zu analysierenden Nachricht voneinander unabhängig sind.