# Maschinelles Lernen II

2. Übung

Prof. Tobias Scheffer
Dr. Niels Landwehr
Christoph Sawade

Sommer 2010

Ausgabe am: 05.05.10
Besprechung am: 12.05.10

**Aufgabe 1** *Bedingte Unabhängigkeit*

Wir betrachten die folgende Domäne: eine faire Münze wird zwei Mal geworfen; die Ergebnisse der Würfe sind repräsentiert durch die Zufallsvariablen $X, Y \in \{0, 1\}$. Wir definieren eine dritte Zufallsvariable $Z$ durch $Z = xor(X, Y)$, d.h. $Z$ hat den Wert Eins, falls genau eine der beiden Variablen $X, Y$ den Wert Eins hat.

1. Geben Sie die gemeinsame Verteilung $p(X, Y, Z)$ über die Variablen $X, Y, Z$ an.
2. (a) Sind $X, Y, Z$ paarweise unabhängig, dh. gilt $p(X, Y) = p(X)p(Y)$, $p(X, Z) = p(X)p(Z)$, und $p(Y, Z) = p(Y)p(Z)$?
   (b) Sind $X, Y, Z$ unabhängig, dh. gilt $p(X, Y, Z) = p(X)p(Y)p(Z)$?
   (c) Geben Sie alle Unabhängigkeiten an, die in der Verteilung $p(X, Y, Z)$ gelten, als Menge
   $$I(p) := \{(A \perp B|C) \mid p(A|B, C) = p(A|C)\}$$
   wobei $A, B, C$ beliebige Teilmengen von $\{X, Y, Z\}$ sind.
3. Geben Sie ein Bayessches Netz für die Verteilung $p(X, Y, Z)$ an. Gilt $I(G) = I(p)$?

**Aufgabe 2** *Repräsentation von bedingten Verteilungen*

In der 1. Übung wurde ein modifiziertes Naïve Bayes Modell diskutiert, dass die Verteilung über Merkmale und die Klasse durch

$$p(\mathbf{x}, y) = \prod_{i=1}^{n} p(x_i)p(y|x_1, ..., x_n)$$

modelliert, wobei

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{pmatrix}$$

der Merkmalsvektor ist.

1. Nehmen Sie $n = 3$ an, und die bedingte Verteilung

| $p(y = 1 \mid x_1, x_2, x_3)$ | $x_1$ | $x_2$ | $x_3$ |
|---|---|---|---|
| 1/2 | 0 | 0 | 0 |
| 2/3 | 1 | 0 | 0 |
| 3/4 | 0 | 1 | 0 |
| 6/7 | 1 | 1 | 0 |
| 4/5 | 0 | 0 | 1 |
| 8/9 | 1 | 0 | 1 |
| 12/13 | 0 | 1 | 1 |
| 24/25 | 1 | 1 | 1 |

   Bekanntlich hat dieses Modell das Problem, dass die Verteilung $p(y \mid x_1, ..., x_n)$ exponentiell viele Parameter enthält. Es ist allerdings auch nicht immer nötig, eine solche bedingte Verteilung explizit als Tabelle zu repräsentieren. Die oben dargestellte Verteilung zeigt gewisse Regelmässigkeiten, und kann viel einfacher mit nur drei Parametern charakterisiert werden.

   Geben Sie eine Repräsentation der Verteilung an, die mit drei Parametern auskommt.

   *Hinweis:* Die Sigmoid-Funktion, gegeben durch

$$\sigma(z) = \frac{1}{1 + exp(-z)}$$

   könnte hier hilfreich sein, ebenso wie lineare Modelle.

2. Nehmen wir an, wir möchten eine Verteilung $p(y \mid x_1, ..., x_n)$ repräsentieren, die nicht die Regelmässigkeiten der oben gegeben Verteilung zeigt. Wir wollen trotzdem nur wenige Parameter im Modell haben, und sind dafür bereit, die Verteilung nur approximativ zu repräsentieren. Wie könnte man in diesem Fall vorgehen?

   *Hinweis:* Maschinelles Lernen ☺.

**Aufgabe 3** *Azyklische Graphen*

Beweisen Sie den folgenden Satz aus der Graphentheorie:

Ein Graph $G$ ist azyklisch genau dann wenn es eine Ordnung $\leq_G$ auf den Knoten von $G$ gibt, so dass für alle $X, X' \in G$ gilt: $X \to X' \Longrightarrow X \leq_G X'$.

Hierbei bezeichnet $X \to X'$, dass es eine gerichtete Kante von Knoten $X$ zu Knoten $X'$ gibt.

*Hinweis:* Es bietet sich an die eine Richtung per Widerspruch und die Gegenrichtung mit Hilfe einer Konstruktionsvorschrift zu beweisen.