

# Maschinelles Lernen und Data Mining

## 2. Übung

Prof. Tobias Scheffer  
Steffen Bickel

WS06/07

Ausgabe am: 30.10.06  
Besprechung am: 03.11.06

### Aufgabe 1 (1/4 Punkt):

Nehmen wir an, dass  $T1$  ein Entscheidungsbaum ist, und  $T2$  ein Entscheidungsbaum, der dadurch entsteht, dass statt eines Blattknotens ein weiterer Split-Knoten eingefügt wird. Besteht zwischen  $T1$  und  $T2$  eine Allgemeinheitsbeziehung, d.h. gilt  $T1 \geq_g T2$  oder  $T2 \geq_g T1$ ? (Beweis oder Gegenbeispiel)

### Aufgabe 2 (2/4 Punkt):

Vergleichen Sie die Lernergebnisse von ID3 mit dem Versionsraumansatz.

- (a) Welchen Baum lernt ID3 (mit Information Gain, nicht mit Gain Ratio), wenn man die 4 Beispiele (SvenSegelt) aus der Versionsraumvorlesung als Eingabe nutzt?
- (b) Ist der gelernte Baum zu einem der Mitglieder des Versionsraums aus der Versionsraumvorlesung äquivalent, oder liegt er außerhalb?
- (c) Könnte man Entscheidungsbäume als Hypothesensprache für das Versionsraumverfahren (candidate elimination) nutzen? (Vorteile? Probleme?)
- (d) Nehmen wir an, für die SvenSegelt-Lernaufgabe sind die Anzahl der möglichen Werte für die 6 Attribute (2,2,2,2,2,2). Wie viele mögliche semantisch unterschiedliche Entscheidungsbäume gibt es dann, die alle mit der Trainingsmenge konsistent sind? Woher weiß ID3, welcher der Bäume der "Richtige" ist?

### Aufgabe 3 (1/4 Punkt):

Werfen Sie einen Blick auf die folgende Trainingsmenge ( $y$  und  $z$  sind diskrete Attribute); simulieren Sie C4.5 von Hand. Wie sieht die Berechnungsspur von C4.5 aus und welcher Baum wird gelernt?

$x$	$y$	$z$	$c$
0.1	2	0	A
0.1	3	0	A
0.2	3	1	A
0.2	4	0	B
0.3	2	0	A
0.35	2	1	B
0.3	3	1	B
0.3	4	1	B