

# Maschinelles Lernen: Symbolische Ansätze

Übungsblatt für den 05.12.2006

## Aufgabe 1

Gegeben sei folgende Beispielmenge:

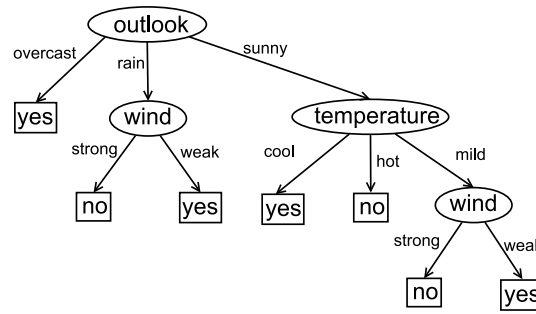
Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	?	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No
D15	Sunny	Mild	Normal	Weak	No

Erzeugen Sie einen Entscheidungsbaum mittels des Verfahrens ID3 (TDIDT mit Maß Gain). Berechnen Sie den Gain des Attributes Outlook, indem Sie Anteile des Beispiels in die Teilbäume propagieren (siehe letzter Punkt auf Folie 34).

Anmerkung: ? steht hier für einen unbekanntem/fehlenden Attributwert.

## Aufgabe 2

Gegeben sei der Entscheidungsbaum aus der Übung vom 28.11.06



und die folgende Pruning-Menge (Validierungsmenge):

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D16	Sunny	Mild	High	Strong	No
D17	Rain	Hot	Normal	Weak	Yes
D18	Overcast	Cool	High	Strong	No
D19	Overcast	Mild	Normal	Strong	Yes
D20	Sunny	Cool	High	Strong	No

Wenden Sie Reduced-Error Pruning (Folie 28) auf den Entscheidungsbaum an. Benutzen Sie als Evaluierungsmaß die Anzahl der korrekt klassifizierten Beispiele der Pruning-Menge.

## Aufgabe 3

Wiederholen Sie die vorherige Aufgabe mit Regel-Post-Pruning (Folie 31). Gehen Sie wie folgt vor:

1. Konvertieren Sie den Baum in eine Regelmenge.
2. Entfernen Sie eine Bedingung einer beliebigen Regel, so daß die Anzahl der korrekt klassifizierten Beispiele
  - a) dieser Regel
  - b) der Regelmenge
 maximiert wird.

3. Wiederholen Sie Aufgabe 3.2 mit den entstandenen Regelmengen (jeweils einmal für die einzelne Regel und für die Regelmenge).