

Knowledge Discovery

Übungsblatt 3

Sommersemester 2004

Aufgabe 1:

Geben Sie für die folgenden booleschen Funktionen einen Entscheidungsbaum an:

- $A \wedge \neg B$
- $A \vee [B \wedge C]$
- $A \text{ XOR } B$
- $[A \wedge B] \vee [C \wedge D]$

Existiert Ihrer Meinung nach eine ähnliche (oder komplexere) Funktion für Aufgabe 2 des letzten Blattes.

Aufgabe 2:

- Geben Sie ein Prinzip an, mit dem man eine aussagenlogische Beschreibung einer (im folgenden als positiv bezeichneten) Klasse aus dem Entscheidungsbaum erzeugen kann.
- Machen Sie das für den Entscheidungsbaum des letzten Übungsblattes für die gesunden Patienten.
- Ein Entscheidungsbaum E2 entsteht durch einen zusätzlichen Split aus dem Entscheidungsbaum E1. Ist dann die zu E1 gehörende Formel allgemeiner als die zu E2 gehörende? (Beweis oder Gegenbeispiel)

Aufgabe 3:

- Berechnen sie für folgenden Datensatz die in der Vorlesung im Detail vorgestellten informationstheoretischen Maße und interpretieren Sie kurz das Ergebnis!

ID	Attribut I	Attribut II	Klasse
1	5	xxx	0
2	42	xxx	1
3	5	zzz	0
4	5	xxx	0
5	42	xxx	1
6	42	zzz	0
7	42	xxx	1
8	42	xxx	1

- Wäre es möglich, den Mittelwert für Attribut I zu berechnen? Wie sinnvoll ist dies?

Aufgabe 4: Begriffe aus dem Bereich des Maschinellen Lernens

- a) Maschinelle Lernverfahren zeichnen sich durch die Eigenschaft des inkrementellen bzw. des nicht-inkrementellen Lernen aus. Was charakterisiert Verfahren mit dieser Eigenschaft?
- b) Was unterscheidet induktives Schließen von deduktivem Schließen?
- c) Induktive Hypothesen werden durch die Begriffe "Vollständigkeit" und "Konsistenz" gekennzeichnet. Beschreiben Sie diese beiden Begriffe mit Hilfe einer Grafik.