

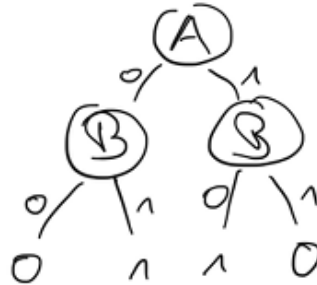
CAL2

Carsten Gips (FH Bielefeld)

Unless otherwise noted, this work is licensed under CC BY-SA 4.0.

Entscheidungsbäume: Klassifikation

A	B	$A \text{ xor } B$
0	0	0
0	1	1
1	0	1
1	1	0



- Attribute als Knoten im Baum
- Ausprägungen als Test (Ausgang, Verzweigung)
- Klasse (Funktionswert) als Blatt

Definition Entscheidungsbaum

- Erinnerung: **Merkmalsvektor** für Objekt v :

$$\mathbf{x}(v) = (x_1, x_2, \dots, x_n)$$

- n Merkmale (Attribute)
- Attribut x_t hat m_t mögliche Ausprägungen
- Ausprägung von v bzgl. x_t : $x_t(v) = i$ (mit $i = 1 \dots m_t$)

- **Alphabet** für Baum:

$$\{x_t | t = 1, \dots, n\} \cup \{\kappa | \kappa = *, A, B, C, \dots\} \cup \{(\,,\,)\}$$

- **Entscheidungsbaum** α :

$$\alpha = \begin{cases} \kappa & \text{Terminalsymbole: } \kappa = *, A, B, \dots \\ x_t(\alpha_1, \alpha_2, \dots, \alpha_{m_t}) & x_t \text{ Testattribut mit } m_t \text{ Ausprägungen} \end{cases}$$

Induktion von Entscheidungsbäumen: CAL2

1) Anfangsschritt: $\alpha^{(0)} = *$ (totales Unwissen)

2) n -ter Lernschritt: Objekt v mit Klasse k , Baum $\alpha^{(n-1)}$ gibt κ aus

- $\kappa = *$: ersetze $*$ durch k
- $\kappa = k$: keine Aktion nötig
- $\kappa \neq k$: Fehler
 - Ersetze κ mit neuem Test: $\kappa \leftarrow x_{t+1}(*, \dots, *, k, *, \dots, *)$
 - x_{t+1} : nächstes Attribut, auf dem aktuellen Pfad noch nicht verwendet
 - Symbol k an Position i wenn $x_{t+1}(v) = i$

Beispiel mit CAL2

x_1	x_2	x_3	k
0	0	1	A
1	0	0	A
0	1	4	B
1	1	2	B
0	0	3	A

Beispiel mit CAL2

x_1	x_2	x_3	k
0	0	1	A
1	0	0	A
0	1	4	B
1	1	2	B
0	0	3	A

Tafelbeispiel CAL2

Ergebnis: $x_1(x_2(A, B), x_2(A, B))$

- Nur für diskrete Merkmale und disjunkte Klassen
- Zyklischer Durchlauf durch Trainingsmenge
- Abbruch:
 - Alle Trainingsobjekte richtig klassifiziert
=> Kein Fehler in einem kompletten Durchlauf
 - (Differenzierung nötig, aber alle Merkmale verbraucht)
 - (Lernschrittzahl überschritten)

- Darstellung der Hypothese als Entscheidungsbaum
- CAL2: diskrete Attribute, disjunkte Klassen

LICENSE



Unless otherwise noted, this work is licensed under CC BY-SA 4.0.