

Blatt 4: Entscheidungsbäume

Aufgabe 1: Entscheidungsbäume mit CAL3 und ID3

Gesamtverteilung: 4 * O, 3 * M

→ Entropie $\approx 0,99$ Bit

Gain:

- Alter = 0,02
- Einkommen = 0,13
- Bildung = 0,31

→ Attribut mit höchstem Gain = Bildung

CAL3:

Bildung	Beispiele	Mehrheit
Master	2	O
Bachelor	2 (1 O, 1 M)	Gleichstand = globale Mehrheit = O
Abitur	3	M

Baum:

Bildung?

└─ Master → O

└─ Bachelor → O

└─ Abitur → M

ID3:

Master → beide O → O

Bachelor → gemischt → split nach Alter ($\geq 35 \rightarrow M$, $< 35 \rightarrow O$)

Abitur → gemischt → split nach Einkommen (hoch → O, niedrig → M)

Bildung	Beispiele	Mehrheit
Master	Beide O	O
Bachelor	gemischt	Split nach Alter
Abitur	gemischt	Split nach Einkommen

Baum:

Bildung?

└─ Master → O

└─ Bachelor →

| Alter?

| └─ $\geq 35 \rightarrow M$

| └─ $< 35 \rightarrow O$

└─ Abitur →

Einkommen?

└─ hoch → O

└─ niedrig → M

Aufgabe 2: Pruning

Gegebener Baum:

$x_3(x_2(x_1(C, A), x_1(B, A)), x_1(x_2(C, B), A))$

Vereinfachter Baum:

$x_3(x_2(C, B), A)$