

	x_1	x_2	x_3	k	ID3 - 7 Klassen, $3 \times M, 4 \times O$
1	\geq	h	a	O	
2	<	n	m	O	Entropie: $-\frac{4}{7} \cdot \log_2 \frac{4}{7} - \frac{3}{7} \log_2 \frac{3}{7} \approx 0,9852$
3	\geq	h	b	M	
4	\geq	n	a	M	
5	\geq	h	m	O	
6	<	h	b	O	
7	<	n	a	M	

Hilfslere
Entropie + Gain

X1) $4 \times \geq, 3 \times <$

\geq : Klassen {O, M, M, O} = $2 \times O, 2 \times M \rightarrow$ Entropie = 1

<: Klassen {O, O, M} = $2 \times O, 1 \times M \rightarrow H = -\frac{2}{3} \log_2 \frac{2}{3} - \frac{1}{3} \log_2 \frac{1}{3} \approx 0,9183$

$$\text{Gain}(S, x_1) = \frac{4}{7} \cdot 1 + \frac{3}{7} \cdot 0,9183 \approx 0,9645$$

$$\text{Gain}(S, x_1) = H(S) - R(S, x_1) \approx 0,9852 - 0,9645 \approx 0,0207$$

X2) $4 \times h, 3 \times n \quad h: \{3 \times O, 1 \times M\} \quad n: \{1 \times O, 2 \times M\}$

$$H(h) = -\frac{3}{4} \log_2 \left(\frac{3}{4}\right) - \frac{1}{4} \log_2 \left(\frac{1}{4}\right) \approx 0,8113$$

$$H(n) = -\frac{1}{3} \log_2 \left(\frac{1}{3}\right) - \frac{2}{3} \log_2 \left(\frac{2}{3}\right) \approx 0,9183$$

$$R(S, x_2) = \frac{4}{7} \cdot 0,8113 + \frac{3}{7} \cdot 0,9183 \approx 0,8572$$

$$\text{Gain}(S, x_2) = H(S) - R(S, x_2) = 0,9852 - 0,8572 \approx 0,128$$

X3) $3 \times a, 2 \times m, 2 \times b \quad a: \{1 \times O, 2 \times M\}, b: \{1 \times M, 1 \times O\} \quad m: \{2 \times O\}$

$$H(a) = -\frac{1}{3} \log_2 \left(\frac{1}{3}\right) - \frac{2}{3} \log_2 \left(\frac{2}{3}\right) = 0,9183$$

$$H(b) = 1 \quad H(m) = 0$$

$$R(S, x_3) = \frac{3}{7} \cdot 0,9183 + \frac{2}{7} \cdot 1 + \frac{2}{7} \cdot 0 \approx 0,6793$$

$$\text{Gain}(S, x_3) = H(S) - R(S, x_3) = 0,9852 - 0,6793 = 0,3059$$

$\Rightarrow \text{Gain}(S, x_3)$ am größten $\rightarrow x_3$ ist Wurzel

X3: a: 1,4,7 $\rightarrow \{O, M, M\}$, Entropie: 0,9183

x_2 innerhalb a:

$$1: h: \{O\} \Rightarrow H(x_3=a, x_2=h) = 0$$

$$4,7: n: \{H, H\} \Rightarrow H(x_3=a, x_2=n) = 0$$

$$\rightarrow R = \frac{1}{3} \cdot 0 + \frac{2}{3} \cdot 0 = 0, \quad \text{Gain} = 0,9183 - 0 = 0,9183$$

x_1 innerhalb a:

$$1,4: \geq: \{O, M\} \Rightarrow H(x_3=a, x_1=\geq) = 1$$

$$7: <: \{M\} \Rightarrow H(x_3=a, x_1=<) = 0$$

$$\rightarrow R = \frac{2}{3} \cdot 1 + \frac{1}{3} \cdot 0 = 0,6667, \quad \text{Gain} = 0,9183 - 0,6667 = 0,2516$$

$\Rightarrow x_2$ trennt vollständig \rightarrow Abitur \rightarrow Einkommen $\begin{cases} \text{nach} \rightarrow O \\ \text{niedrig} \rightarrow M \end{cases}$

X3: m: 2,5 $\rightarrow \{O, O\}$, Entropie = 0 \rightarrow keine weitere Teilung nötig

X3: b: 3,6 $\rightarrow \{M, O\}$, Entropie = 1

x_1 innerhalb b:

$$3: \geq: \{M\} \Rightarrow H(x_3=b, x_1=\geq) = 0$$

$$6: <: \{O\} \Rightarrow H(x_3=b, x_1=<) = 0$$

$$R=0, \quad \text{Gain} = 0,9183 - 0 = 0,9183$$

$\Rightarrow x_1$ trennt vollständig \rightarrow Bachelor \rightarrow Alter $\begin{cases} \geq 35 \rightarrow M \\ < 35 \rightarrow O \end{cases}$

Bildung (x_3)

Prüfen auf Trainingsfehler
 \rightarrow keine Fehler!

