

Decision Trees Exercise

Aufgabe a)

1. Entropie Buys-Computer berechnen

→ 8x No } 20
→ 12x Yes

$$H(\text{Buys-Computer}) = H(8, 12) = -(0,4 \log_2 0,4) - (0,6 \log_2 0,6) \\ = \underline{\underline{0,97}}$$

2. Entropien aller Attribute berechnen

Age	No	Yes	
≤ 30	6x	2x	8
31...40	0x	6x	6
> 40	2x	4x	6

$$E(\text{Buys-Computer, Age}) = P(\leq 30) E(6, 2) + P(31...40) E(0, 6) + P(> 40) E(2, 4) \\ = \frac{8}{20} \cdot 0,81 + \frac{6}{20} \cdot 0 + \frac{6}{20} \cdot 0,92 \\ = \underline{\underline{0,6}}$$

Income

	No	Yes	
Low	3x	4x	7
Medium	3x	5x	8
High	2x	3x	5
			20

$$E(\text{Buys-Computer, Income}) = P(\text{Low}) \cdot E(3, 4) + P(\text{Medium}) \cdot E(3, 5) + P(\text{High}) \cdot E(2, 3) \\ = \frac{7}{20} \cdot 0,99 + \frac{8}{20} \cdot 0,95 + \frac{5}{20} \cdot 0,97 \\ = \underline{\underline{0,96}}$$

Student

	No	Yes	
No	7x	4x	11
Yes	7x	8x	15
			20

$$\begin{aligned} E(\text{Buys_computer}, \text{Student}) &= P(\text{No}) \cdot E(7, 4) + P(\text{Yes}) \cdot E(7, 8) \\ &= \frac{11}{20} \cdot 0,94 + \frac{9}{20} \cdot 0,5 \\ &= 0,74 \end{aligned}$$

Credit Rating

	No	Yes	
Fair	3x	7x	10
Excellent	5x	5x	10
			20

$$\begin{aligned} E(\text{Buys_computer}, \text{Credit Rating}) &= P(\text{Fair}) \cdot E(3, 7) + P(\text{Excellent}) \cdot E(5, 5) \\ &= \frac{10}{20} \cdot 0,88 + \frac{10}{20} \cdot 1 \\ &= 0,94 \end{aligned}$$

3. Information Gain berechnen

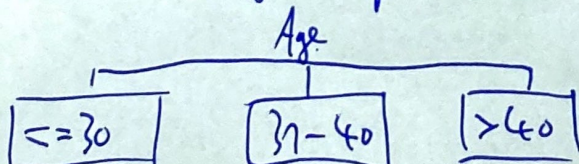
$$\begin{aligned} \text{Age: } H(\text{Buys_computer}) - E(\text{Buys_computer}, \text{Age}) \\ &= 0,97 - 0,6 = 0,37 \end{aligned}$$

$$\text{Income: } 0,97 - 0,97 = 0$$

$$\text{Student: } 0,97 - 0,74 = 0,23$$

$$\text{Credit Rating: } 0,97 - 0,94 = 0,03$$

→ 7. Split bei Attribut "Age" verspricht höchsten Information Gain



Aufgabe b)



WebGraphviz is [Graphviz](#) in the Browser

Enter your graphviz data into the Text Area:

(Your Graphviz data is private and never harvested)

Sample 1

Sample 2

Sample 3

Sample 4

Sample 5

```
digraph Tree {
  node [shape=box, style="filled, rounded", color="black", fontname="helvetica"] ;
  edge [fontname="helvetica"] ;
  0 [label="Age <= 0.5\ngini = 0.48\nsamples = 20\nvalue = [8, 12]\nclass = buy", fi
  1 [label="Student <= 0.5\ngini = 0.375\nsamples = 8\nvalue = [6, 2]\nclass = doesn
  0 -> 1 [labeldistance=2.5, labelangle=45, headlabel="True"] ;
  2 [label="gini = 0.0\nsamples = 6\nvalue = [6, 0]\nclass = doesnt buy", fillcolor=
  1 -> 2 ;
  3 [label="gini = 0.0\nsamples = 2\nvalue = [0, 2]\nclass = buy", fillcolor="#399de
  1 -> 3 ;
  4 [label="Credit_rating <= 0.5\ngini = 0.278\nsamples = 12\nvalue = [2, 10]\nclass
  0 -> 4 [labeldistance=2.5, labelangle=-45, headlabel="False"] ;
  5 [label="gini = 0.0\nsamples = 6\nvalue = [0, 6]\nclass = buy", fillcolor="#399de
  4 -> 5 ;
  6 [label="gini = 0.444\nsamples = 6\nvalue = [2, 4]\nclass = buy", fillcolor="#9cc
  4 -> 6 ;
}
```

Generate Graph!

