

Exercise 3

$$\begin{aligned} 1) \quad N &= 400 \\ A &: (300, 100) \quad (100, 300) \\ B &: (200, 0) \quad (200, 400) \end{aligned}$$

miss classification rate:

$$A: \frac{200}{800} = \underline{\underline{\frac{1}{4}}}$$

$$B: \frac{200}{800} = \underline{\underline{\frac{1}{4}}}$$

gini impurity:

$$\begin{aligned} A: H_1 &= 1 - (p(y=1)^2 + p(y=2)^2) \\ &= 1 - \left(\left(\frac{3}{4} \right)^2 + \left(\frac{1}{4} \right)^2 \right) \\ &= 1 - \frac{9}{16} - \frac{1}{16} \\ &= \frac{6}{16} = \underline{\underline{\frac{3}{8}}} \end{aligned}$$

$$H_2 = \underline{\underline{\frac{3}{8}}}$$

$$H_A = \underline{\underline{\frac{3}{8}}}$$

$$B: H_1 = 1 - (1^2 + 0^2) = \underline{0}$$

$$H_2 = 1 - \left(\left(\frac{2}{6} \right)^2 + \left(\frac{4}{6} \right)^2 \right)$$

$$= 1 - \frac{4}{36} - \frac{16}{36}$$

$$= \frac{16}{36} = \underline{\underline{\frac{4}{9}}}$$

$$H_B = \frac{200}{800} \cdot 0 + \frac{600}{800} \cdot \frac{4}{9}$$

$$= \frac{3}{4} \cdot \frac{4}{9}$$

$$= \underline{\underline{\frac{1}{3}}}$$

Entropy:

$$A: H_A = \frac{3}{4} \cdot \log\left(\frac{3}{4}\right) - \frac{1}{4} \log\left(\frac{1}{4}\right)$$

$$\approx \underline{\underline{0,56}}$$

$$B: H_1 = 0$$

$$H_2 = -\frac{1}{3} \cdot \log\left(\frac{1}{3}\right) - \frac{2}{3} \cdot \log\left(\frac{2}{3}\right)$$

$$\approx \underline{\underline{0,64}}$$

$$H_B = H_2 \cdot \frac{3}{4}$$

$$= 0,64 \cdot 0,75$$

$$= \underline{\underline{0,48}}$$

Answer: The misclassification rate is the same for both splits. But the gini impurity and the entropy of split B are better than those of split A.
That is why we should consider split B.

2) a)

	0	0	0	$0 \cdot 1 + 0 \cdot 1 - 0,5 = -0,5 \Rightarrow 0$
	0	1	1	$0 \cdot 1 + 1 \cdot 1 - 0,5 = 0,5 \Rightarrow 1$
	1	0	1	$1 \cdot 0 + 0 \cdot 0 - 0,5 = -0,5 \Rightarrow 1$
	1	1	1	$1 \cdot 0 + 1 \cdot 0 - 0,5 = -0,5 \Rightarrow 1$

$$\text{weight} = [1_1, 1_2, \dots, 1_n]$$

$$\text{bias} = -0,5$$

b)

	0	0	0	$0 \cdot 1 + 0 \cdot 1 - 2 = -2 \Rightarrow 0$
	0	1	0	$0 \cdot 1 + 1 \cdot 1 - 2 = -1 \Rightarrow 0$
	1	0	0	$1 \cdot 1 + 0 \cdot 1 - 2 = -1 \Rightarrow 0$
	1	1	1	$1 \cdot 1 + 1 \cdot 1 - 2 = 0 \Rightarrow 1$

$$\text{weight} = [1_1, 1_2, \dots, 1_n]$$

$$\text{bias} = -n$$

c)

rot	1	0	0	$1 \cdot (-1) + 0 \cdot (-1) + 0,5 = -0,5$
grün	0	0	1	$0 \cdot (-1) + 0 \cdot (-1) + 0,5 = 0,5$
gelb	0	1	0	$0 \cdot (-1) + 1 \cdot (-1) + 0,5 = -0,5$
blau	1	1	0	$1 \cdot (-1) + 1 \cdot (-1) + 0,5 = -1,5$

$$\text{weights} = [-1, -1]$$

$$\text{bias} = 0,5$$

d)

rot	1	0	1	$1 \cdot (1) + 0 \cdot (-1) - 0,5 = 0,5$
grün	0	0	0	$0 \cdot (1) + 0 \cdot (-1) - 0,5 = -0,5$
gelb	0	1	0	$0 \cdot (1) + 1 \cdot (-1) - 0,5 = -1,5$
blau	1	1	0	$1 \cdot (1) + 1 \cdot (-1) - 0,5 = -0,5$

$$\text{weights} = [1, -1]$$

$$\text{bias} = -0,5$$

3) a)

$$\begin{aligned}\frac{d}{dw} h(f(x_1, w), g(x_2, w)) &= \frac{dh}{dw} \\ &= \frac{dh}{df} \cdot \frac{df}{dw} + \frac{dh}{dg} \cdot \frac{dg}{dw}\end{aligned}$$

b) Chain rule

$$c) \frac{dh}{df} = \frac{d}{df} (f \cdot g) = \underline{\underline{g}}$$

$$\frac{df}{dw} = \frac{d}{dw} (x_1 \cdot w) = \underline{\underline{x_1}}$$

$$\frac{dh}{dg} = \frac{d}{dg} (f \cdot g) = \underline{\underline{f}}$$

$$\frac{dg}{dw} = \frac{d}{dw} (x_2 \cdot w) = \underline{\underline{x_2}}$$

$$\begin{aligned}\frac{dh}{dw} &= g \cdot x_1 + f \cdot x_2 = (x_2 \cdot w) \cdot x_1 + (x_1 \cdot w) \cdot x_2 \\ &= w \cdot (x_2 \cdot x_1 + x_1 \cdot x_2) \\ &= w \cdot (2 \cdot x_1 \cdot x_2) \\ &= \underline{\underline{2 \cdot x_1 \cdot x_2 \cdot w}}\end{aligned}$$

$$d) \frac{dh}{df} = \frac{d}{df} (f^g) = \underline{\underline{g \cdot f^{(g-1)}}}$$

$$\frac{df}{dw} = \frac{d}{dw} (x_1 + w) = \underline{1}$$

$$\frac{dh}{dg} = \frac{d}{dg} (f^g) = \underline{f^g \cdot \ln(f)}$$

$$\frac{dg}{dw} = \frac{d}{dw} (x_2 + w) = \underline{1}$$

$$\frac{dh}{dw} = g \cdot f^{(g-1)} + f^g \cdot \ln(f)$$

$$= (x_2 + w) \cdot (x_1 + w)^{(x_2 + w - 1)} + (x_1 + w)^{(x_2 + w) \cdot \ln(x_1 + w)}$$

$$= (x_2 + w) \cdot (x_1 + w)^{\frac{(x_2 + w)}{(x_1 + w)}} + (x_1 + w)^{(x_2 + w) \cdot \ln(x_1 + w)}$$

$$= \underline{(x_1 + w)^{(x_2 + w)} \cdot \left(\frac{x_2 + w}{x_1 + w} + \ln(x_1 + w) \right)}$$