We consider the follwing data set

$$\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{5} = \{(1,0), (2,1), (3,0), (4,1), (5,1)\}$$

and compute the information gain associated with a split at $\xi = 3.5$. The entropy of the whole dataset is

$$\begin{aligned} H(\mathcal{D}) &= -\left(p_0 \log_2 p_0 + p_1 \log_2 p_1\right), \quad p_0 = \frac{2}{5}, \quad p_1 = \frac{3}{5} && (1) \\ &= -\frac{1}{5}\left(2 + 3\log_2 3 - 5\log_2 5\right) && (2) \end{aligned}$$

The first set after the split is

$$\mathcal{D}_1 = \{(1,0), (2,1), (3,0)\}$$

and its entropy is

$$\begin{aligned} H(\mathcal{D}_1) &= -\left(p_0 \log_2 p_0 + p_1 \log_2 p_1\right), \quad p_0 = \frac{2}{3}, \quad p_1 = \frac{1}{3} && (3) \\ &= -\frac{1}{3}\left(2 - 3\log_2 3\right) && (4) \end{aligned}$$

The second set after the split is

$$\mathcal{D}_2 = \{(4,1), (5,1)\}$$

and its entropy is

$$\begin{aligned} H(\mathcal{D}_2) &= -\left(p_0 \log_2 p_0 + p_1 \log_2 p_1\right), \quad p_0 = 0, \quad p_1 = 1 && (5) \\ &= 0 && (6) \end{aligned}$$

The information gain is

$$\begin{aligned} \Delta H &= 5H(\mathcal{D}) - \left(3H(\mathcal{D}_1) + 2H(\mathcal{D}_2)\right) && (7) \\ &= -\left(2 + 3\log_2 3 - 5\log_2 5\right) + \left(2 - 3\log_2 3\right) && (8) \\ &= -\left(2 + 3\log_2 3 - 5\log_2 5 - 2 + 3\log_2 3\right) && (9) \\ &= -\left(6\log_2 3 - 5\log_2 5\right) && (10) \\ &\approx -\left(6\frac{16}{10} - 5\frac{23}{10}\right) && (11) \\ &= -\frac{96 - 115}{10} && (12) \\ &= \frac{19}{10} = 1.9 && (13) \end{aligned}$$