Pattern Recognition Assignment#4

61821313 Zihao Wang

May 15, 2023

Question1

Here is the decision tree created through information gain

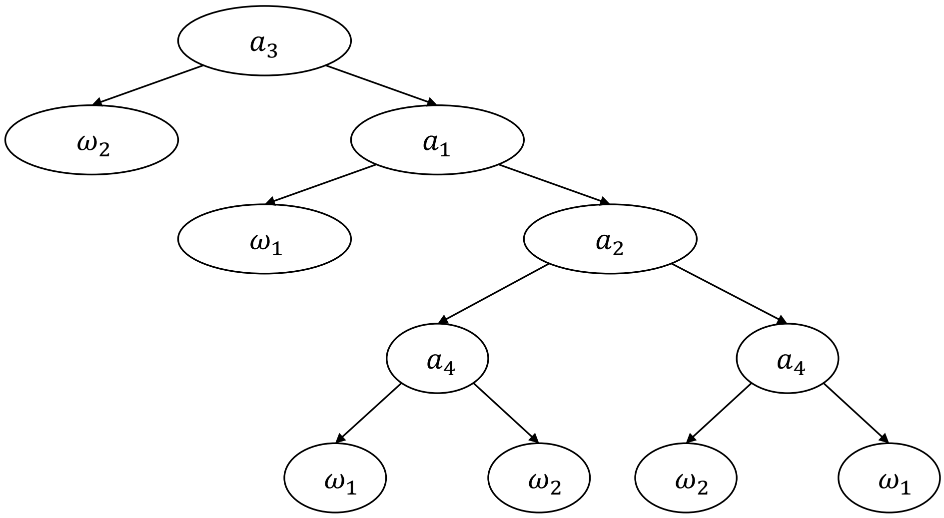

Figure 1: Decision Tree

Where the left child and the right child of a node represent value 0 and 1, respectively

Question2

Calculate the hidden layer

$$net(h_1) = w_{11}^{(1)} x_1 + w_{12}^{(1)} x_2 = 1.6$$

$$net(h_2) = w_{21}^{(1)} x_1 + w_{22}^{(1)} x_2 = 2.8$$

$$h_1 = f(net(h_1)) = 0.8320$$

$$h_2 = f(net(h_2)) = 0.9427$$

Calculate the output layer

$$net(o) = w_{11}^{(2)} h_1 + w_{12}^{(2)} h_2 = 1.1591$$

$$o = f(net(o)) = 0.7612$$

Use the square loss as the loss function

$$\ell = \frac{1}{2}(t - o)^2$$

Calculate the dirivative of the net activation of output layer

$$\frac{\partial \ell}{\partial o} = o - t = -0.2388$$

$$\frac{\partial o}{\partial net(o)} = o(1 - o) = 0.1818$$

$$\frac{\partial \ell}{\partial net(o)} = \frac{\partial \ell}{\partial o} \frac{\partial o}{\partial net(o)} = -0.0434$$

Calculate the dirivative of the hidden-to-output layer weights and the hidden layer

$$\frac{\partial net(o)}{\partial w_{11}^{(2)}} = h_1 = 0.8320 \qquad \frac{\partial \ell}{\partial w_{11}^{(2)}} = \frac{\partial \ell}{\partial net(o)} \frac{\partial net(o)}{\partial w_{11}^{(2)}} = -0.0361$$

$$\frac{\partial net(o)}{\partial w_{12}^{(2)}} = h_2 = 0.9427 \qquad \frac{\partial \ell}{\partial w_{12}^{(2)}} = \frac{\partial \ell}{\partial net(o)} \frac{\partial net(o)}{\partial w_{12}^{(2)}} = -0.0409$$

$$\frac{\partial net(o)}{\partial h_1} = w_{11}^{(2)} = 0.6 \qquad \frac{\partial \ell}{\partial h_1} = \frac{\partial \ell}{\partial net(o)} \frac{\partial net(o)}{\partial h_1} = -0.0260$$

$$\frac{\partial net(o)}{\partial h_2} = w_{12}^{(2)} = 0.7 \qquad \frac{\partial \ell}{\partial h_2} = \frac{\partial \ell}{\partial net(o)} \frac{\partial net(o)}{\partial h_2} = -0.0304$$

Calculate the dirivative of the net activation of hidden layer

$$\frac{\partial h_1}{\partial net(h_1)} = h_1(1 - h_1) = 0.1398$$

$$\frac{\partial h_2}{\partial net(h_2)} = h_2(1 - h_2) = 0.0540$$

$$\frac{\partial \ell}{\partial net(h_1)} = \frac{\partial \ell}{\partial h_1} \frac{\partial h_1}{\partial net(h_1)} = -3.6348 \times 10^{-3}$$

$$\frac{\partial \ell}{\partial net(h_2)} = \frac{\partial \ell}{\partial h_2} \frac{\partial h_2}{\partial net(h_2)} = -1.6416 \times 10^{-3}$$

Calculate the dirivative of the input-to-hidden layer weights

$$\frac{\partial net(h_1)}{\partial w_{11}^{(1)}} = x_1 = 2 \qquad \frac{\partial \ell}{\partial w_{11}^{(1)}} = \frac{\partial \ell}{\partial net(h_1)} \frac{\partial net(h_1)}{\partial w_{11}^{(1)}} = -7.2696 \times 10^{-3}$$

$$\frac{\partial net(h_1)}{\partial w_{12}^{(1)}} = x_2 = 4 \qquad \frac{\partial \ell}{\partial w_{12}^{(1)}} = \frac{\partial \ell}{\partial net(h_1)} \frac{\partial net(h_1)}{\partial w_{12}^{(1)}} = -1.4539 \times 10^{-2}$$

$$\frac{\partial net(h_2)}{\partial w_{21}^{(1)}} = x_1 = 2 \qquad \frac{\partial \ell}{\partial w_{11}^{(1)}} = \frac{\partial \ell}{\partial net(h_2)} \frac{\partial net(h_2)}{\partial w_{21}^{(1)}} = -3.2832 \times 10^{-3}$$

$$\frac{\partial net(h_2)}{\partial w_{22}^{(1)}} = x_2 = 4 \qquad \frac{\partial \ell}{\partial w_{22}^{(1)}} = \frac{\partial \ell}{\partial net(h_2)} \frac{\partial net(h_2)}{\partial w_{22}^{(1)}} = -6.5664 \times 10^{-3}$$

Update the layer weights $\boldsymbol{w}$ through the dirivative

$$\boldsymbol{w} \leftarrow \boldsymbol{w} - \eta \frac{\partial \ell}{\partial \boldsymbol{w}}$$

The resulting model $\boldsymbol{w}$ after training is

$$\boldsymbol{w} = \{0.2073,\ 0.3145,\ 0.4033,\ 0.5066,\ 0.6361,\ 0.7409\}$$