Pattern Recognition Assignment#4

May 10, 2023

Q1. Consider a binary classification problem for 8 samples. Each sample has 4 binary attributes $A_1,\ A_2,\ A_3,\ A_4$, denoted as $a_1a_2a_3a_4$, where $a_i \in \{0, 1\}$. The detail of the dataset is listed in the table below. Please use the information gain as the measure to create an unpruned decision tree for this dataset.

sample	value	category
x_1	0110	w_1
x_2	1010	w_1
x_3	0011	w_1
x_4	1111	w_1
x_5	1011	w_2
x_6	0000	w_2
x_7	0100	w_2
x_8	1110	w_2

Q2. Consider a three-layer feedforward neural network with two input neurons, two hidden neurons, and one output neuron. The architecture is shown below. The activation function of the hidden layer and the output layer is the Sigmoid function $f(x) = \frac{1}{1+e^{-x}}$. Given the initial model $\mathbf{w} = \left\{w_{11}^{(1)}, w_{12}^{(1)}, w_{21}^{(1)}, w_{22}^{(1)}, w_{11}^{(2)}, w_{12}^{(2)}\right\} = \{0.2,\ 0.3, 0.4, 0.5, 0.6, 0.7\}$, where $\left\{w_{11}^{(1)}, w_{12}^{(1)}, w_{21}^{(1)}, w_{22}^{(1)}\right\}$ correspond to the input-to-hidden layer weights and $\left\{w_{11}^{(2)}, w_{12}^{(2)}\right\}$ correspond to the hidden-to-output layer weights, one can yield the optimization using the stochastic backpropagation algorithm. Considering the training sample $\mathbf{x} = [2, 4]^\top$ with the target $t = 1$, what is the resulting model $\mathbf{w}$ after training the neural network with the sample (learning rate $\eta = 1$)? (Tips: you can round the results to the fourth decimal place).

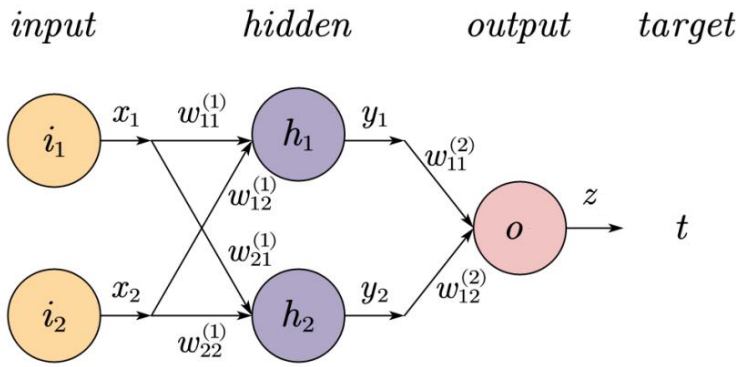

Figure 1: The architecture of a three-layer feedforward neural network.