Intelligent Systems: Reasoning and Recognition

James L. Crowley

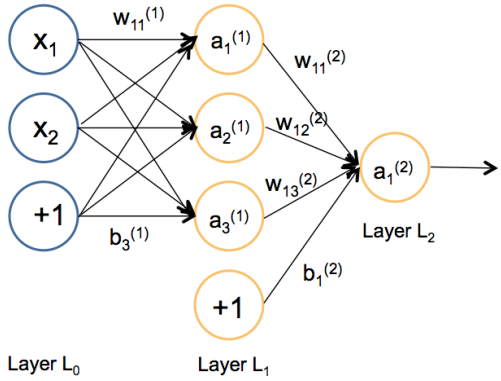
MoSIG M1 Winter Semester 2018-219

Exercise 5 21 February 2019

Artificial Neural Networks.

The following by may be performed by computer program or by hand. The recommended method is to write an interactive Python program.

Assume that you have a 2 layer network with 3 hidden units, and one output, using a sigmoid activation function.



The weights for the first layer have been initialized to the following (random) parameters:

$$W_{ji}^{(1)} = \begin{pmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \end{pmatrix}, \; b_j^1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \; W_{ji}^{(2)} = \begin{pmatrix} 1 & -1 & 0 \end{pmatrix} \text{ and } b_j^2 = (1)$$

You have the following 4 samples of training data.

$$\{\vec{X}_m\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \; \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \; \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \; \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \text{ with } \{y_m\} = \{1,0,1,0\}$$

a) Calculate the output of each unit $a_{ji}^{(l)}$ for the first training sample $\vec{X}_1$.

b) Calculate the error term $\delta_{j,1}^{(l)}$ for each unit by back-propagation for the first training sample $\vec{X}_1$.

c) Calculate the corrections $\Delta w_{ji,1}^{(l)}$ and $\Delta b_{j,1}^{(l)}$ for the first training sample $\vec{X}_1$.

d) Apply the correction to the network parameters using a learning rate of η=0.5 and repeat steps a, b and c for training samples m=2, 3 and 4. What are the resulting network parameters?

e) Determine $\Delta w_{ji,m}^{(l)} = a_i^{(l-1)} \delta_{j,m}^{(l)}$ and $\Delta b_{j,m}^{(l)} = \delta_{j,m}^{(l)}$ for all 4 training sample $\{\vec{X}_m\}$ and $\{y_m\}$ without applying the correction to the weights. Compute the average of the correction factors and then update the weights with a learning rate of η=0.5 using the average. What are the resulting network parameters?