HW: Chapter 11

Multi-layer Perceptrons (Neural Networks) Assume bias is always 1?

1. Suppose we have a single two-input perceptron with weights:
   $w_b = 0.5$, $w_1 = 0.7$ and $w_2 = 0.22$
   and inputs:
   $I_1 = -4$, $I_2 = 12$

   a) Calculate the output for a threshold function of T = 0

   $1(0.5) + -4(0.7) + 12(0.22) = 0.14$

   $0.14 > 0$, output 1

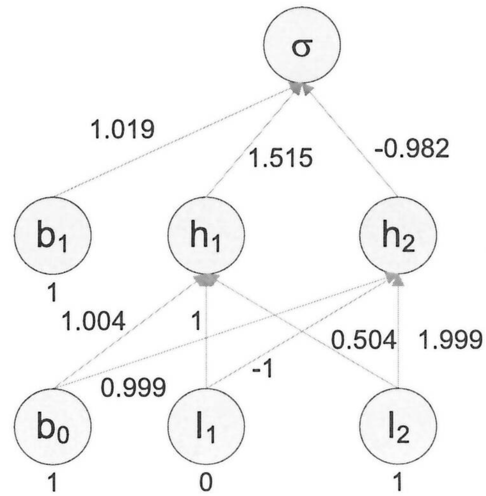
   b) Calculate the output for a threshold function of T = 0.5

   $0.14 < 0.5$, output 0

   c) Calculate the output when the sigmoid function is used

   $\frac{1}{1+e^{-0.14}} = 0.53$, output 1

2. The neural network given below is the final one derived in the tutorial.




error before: 0.024
assume t is 1 again

$\sigma = \sum_i w_i x_i + bias$

$y_{h_i} = \frac{1}{1+e^{-\sigma}}$

- Feedforward; then calculate total Error in the network. Has it been reduced?

$h_1: \sigma = 1.508$
$h_1 = 0.819$

$h_2: \sigma = 2.998$
$h_2 = 0.952$

$y: \sigma = 1.32$
$y = 0.790$

$\text{Error} = \frac{1}{2}(t-y)^2$
$= \frac{1}{2}(1-0.790)^2 = \underline{0.022}$, reduced

3. In neural network training, summing the error over all training examples and then adjusting weights is called *batch* learning, while performing these steps after each training example is called *online* learning. What might be some differences between these approaches with respect to: epoch?
   - resistance to local minima?
   - convergence time?

Batch learning could be more resistant to local minima due to a slower change in error, while online learning updates more frequently and could show a local minima that appears to be global due to a more frequent change in error.

Convergence time would be slower for batch I believe as number of updates required to reach convergence increases with training set size. Online learning has more frequent updates, so convergence would be faster as change is more rapid.