CSC 4510 – Review Questions for Neural Nets

1. One of the motivations for neural nets is that they can learn non-linear functions. We saw, however, that non-linear functions can be learned using logistic regression, by introducing quadratic or higher-order terms. Why is this issue relevant then?

2. What is the “one learning algorithm” hypothesis and what is some experimental evidence that appears to support it? How is it relevant to the study of neural nets?

3. Discuss some similarities and differences between artificial neural networks and the brain.

4. What are some successful applications of neural nets? What are some limitations and shortcomings?

5. Suppose we have a single perceptron as shown below. Use the perceptron update formula to update the weights.

\[ w_i(p+1) = w_i(p) + \alpha * x_i(p) * err(p) \]

6. Show examples of perceptrons that have been trained to compute the AND, OR, and NOT functions. Why can the perceptron learn only linearly separable functions?

7. We experimented in class training neural nets using AI Space. One of the examples looked like this:

% Predicting whether someone will take a class. Binary values, 1 for true, 0 for false.
T:required, have-prereqs, workload, take-class
1, 0, 0, 0
1, 0, 1, 0
1, 1, 0, 0
1, 1, 1, 0
0, 0, 0, 1
0, 0, 1, 1
0, 1, 0, 1
0, 1, 1, 1
% Inputs same as previous 8, but outcomes reversed
1, 0, 0, 1
1, 0, 1, 1
1, 1, 0, 1
1, 1, 1, 1
0, 0, 0, 0
0, 0, 1, 0
0, 1, 0, 0
0, 1, 1, 0

What would you expect the outcome to be when training the network with these examples? Did the network converge? Why or why not?
8. Construct a fully connected multilayer artificial neural network with an input layer of six neurons, a hidden layer of four neurons and an output layer of two neurons. What is a hidden layer for and what does it “hide”?

9. In our experiments with AI Space we saw that we may have non-binary attributes and that these attributes can be ordered or non-ordered. How does the fact that an attribute is ordered affect the training of a neural net?

10. What is the difference between training error and test error? In some of our experiments with AI Space the training and test errors seemed to coincide, whereas in one of the experiments we saw that as the training error went down, the test error went up. What could be the reason for this? Discuss.

11. Discuss your main observations when you created a neural net in AI Space with our class data example (apples or oranges).

12. With perceptrons we used a threshold on the weighted sum of inputs to determine whether the neuron would “fire” or not. For the backpropagation algorithm, however, we switched to applying the sigmoid function instead. Why?

13. Define the following terms used in backpropagation:
   - epoch
   - training error
   - test error
   - hidden unit

14. Use the multi-layer network from Fig 6.10 in the text with some different inputs (eg 0,1), assuming all the other values are the same and that the function being learned is the XOR function. Rework this example to ensure you understand all the steps.