CS 440 - Homework 7

Matthew Liu - mliu56

1. Conceptual Questions

1. What is the example space, $\boldsymbol{X}$, for this problem?
   Since the input can be 0, 0.5, 1, $\boldsymbol{X}$ is 3-dimensional real space R^3

2. Which model do you expect to be more expressive: a binary classifier perceptron or a binary classifier multi-layer perceptron with one hidden layer? Why?
   A binary classifier multi-layer perceptron with one hidden layer will be more expressive, it has more VC dimensions then the former and thus will be the richer model and be able to represent more things.

3. Which one of the models described in question 2 do you think is more likely to overfit, and why?
   Even though the multi-layer perceptron with one hidden layer will be more expressive, it is also more likely to overfit when you do not have enough data.

4. Do you expect your perceptron to converge (assuming a constant learning rate)? Why or why not?
   Yes, because it seems there is plenty of data available to properly train the perceptron, and the data can be properly separated into positive and negative examples by a hyperplane

2. Implementing a Perceptron

Reference: https://en.wikipedia.org/wiki/Perceptron

1. Try adjusting each setting listed above, and report which combination (learning rate, training example ordering, and weight initialization) worked best. If you used random initial weights, specify what distribution the weights came from (uniform, Gaussian, or another distribution).

2. How many epochs were necessary to train your perceptron?

3. What accuracy did you achieve on the training set?

4. What accuracy did you achieve on the test set?

2. Using a Multilayer Perceptron

1. Which combinations of hyperparameters did you try, and which combination worked best?

2. How many epochs were necessary to train your multilayer perceptron?

3. What accuracy did you achieve on the training set?

4. What accuracy did you achieve on the test set?

5. How did the results obtained using the multilayer perceptron compare to those obtained using the single-layer perceptron in part 2? Explain the differences.