A6 Report

Name(s): Olivia Fang

Assignment 6: Perceptron Classification and Training

CSE 415 Introduction to Artificial Intelligence, Winter 2024, University of Washington

Please answer each question using text in Blue, so your answers stand out from the questions.

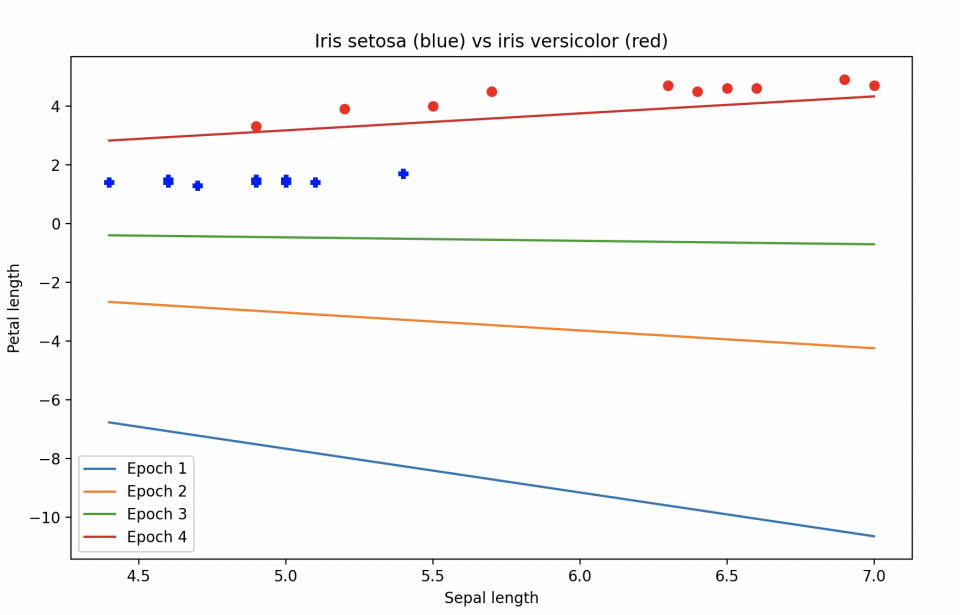
Note: If not otherwise specified, use the default parameters present in the starter code to answer the questions.

Q1. How many epochs were required to train your perceptron on the 2-class Iris data having 2 features? What was the performance of your perceptron on the test data?

4 epochs were required to train my perceptron on the Iris data.

The performance on the test data is 2 errors out of 80 times.

Q2. Include a graphic produced using matplotlib that shows both the training data points (in separate colors) and the “separating” lines implied by the weights at the end of each training epoch.” (Reduce the graphic as necessary to make it fit here without taking up more than half the page.)



Q3. In the above plot, was there any thrashing (oscillation in the separator, such as flipping slope back and forth between positive and negative values, or having its y intercept jumping up and down as epochs proceed? How would you describe the progress of the learning, on the basis of the plot?

I can see that the separator initially started with low y-intercept and negative slope. Throughout the learning process, it figured out to increase the intercept by a bit each time, and increase the slope, so the separator eventually ended up with an intercept between the blue and red data points, with a

positive slope that follows the trend of the data. So there is not much trashing, the change is consistent.

Q4. After plotting the ring data, describe its distribution in words.

Before remapping, the ring data's distribution is like its name, in rings, with the red data outside and the blue inside the ring formed by the red data. So there is no way a straight line could separate them into different clusters.

Q5. Describe the sequence of separators obtained when training your perceptron for 25 epochs using the ring data. Is there any thrashing? To what extent did it achieve convergence? And finally, do you think if the model is run for more epochs it will eventually fully converge?

There is definitely thrashing going on the whole time it is being trained. The separators go back and forth between having a high intercept and negative slope, and having a low intercept with positive slope. So these separators cross in the middle where the middle of the ring is. But it cannot converge, since it cannot figure out a way to separate the data points with a straight line.

As I said in the previous question, it is not possible to separate the two rings of data into two clusters with one straight line. So even if we run the model infinite times, it will still not fully converge.

Q6. After you have re-mapped the ring data with the provided non-linear mapping function, plot the data and describe the distribution.

After remapping, the distribution is more linear, with the points lining up, rather than circling up.

Q7. After training your perceptron on the re-mapped ring data, did it achieve convergence, and if so, how many epochs were used?

It took 11 epochs to fully converge.

From the plot, I can see that there is some thrashing going on. It first started rather close to the points, with a slope close to how the data is distributed, but after a few minor adjustments, the slope suddenly became very steep and negative. Then it goes back to the original flat slope, then it tries with a positive steep slope, but with a more accurate intercept. Then it probability decided that very positive or slopes are both bad, so it went back with the flat slope, and adjusted slightly each time until it converged.

Q8. What do these results suggest about the power of perceptrons to classify data that may consist of clusters that cannot be separated by a linear manifold (such as a line or plane)?

The perceptron model converged after we remapped the points into a linear relationship. This suggests that perceptrons cannot classify data well if the data consists of clusters that cannot be separated by a linear manifold.