A6 Report

Name: Shinji Yamashita

Assignment 6: Perceptron Classification and Training

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 for the weights to converge on my training data. For the test data, there were only two errors out of 80, indicating a 97.5 percent accuracy.

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.)

A diagram of different colored lines

Description automatically generated

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?

No, there wasn’t really any oscillation as the slope gradually became from negative to positive consistently without the sign of the slope switching back and forth.

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

There seems to be no linear separability as the blue points are gathered in a ring like distribution in the middle while the red points are gathered in a ring like distribution that encloses the smaller ring.

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 was trashing as the slope often oscillated between positive and negative between the epochs. At 25 epochs, there was a balance of four blue data points on both sides of the line. The rest on both sides were red points. I believe that the model will not fully converge even after running infinitely many epochs, since at epoch 25, the slope was very similar to the slope in the middle of the epochs so it would just keep flipping back and forth, never fully converging.

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

The data points are linearly separable, as the blue and red points seem to be gathered in two separate clusters according to their color where the blue points have a smaller r, whereas the red points have a bigger r in general.

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

It fully converged in 11 epochs.

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 power of perceptron’s is limited by the assumption that the data must be linearly separable.