Part A Report

Name: Khoa Tran (replace with your own name and delete this instruction!)

Assignment 6: Perceptron Classification and Training

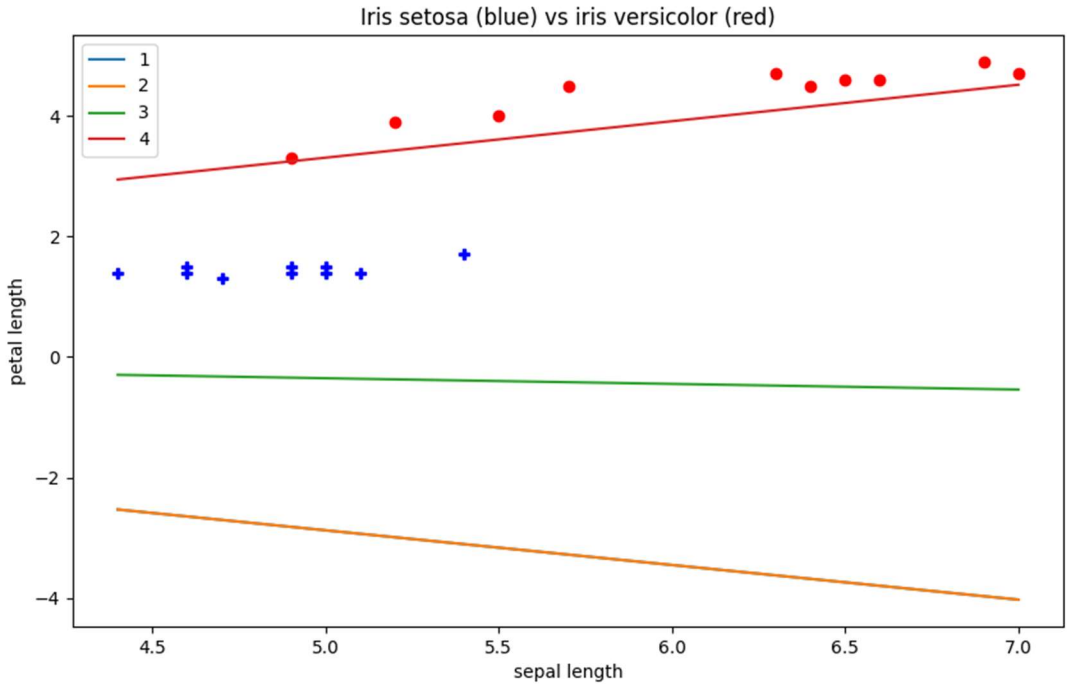
CSE 415 Introduction to Artificial Intelligence, Winter 2021, University of Washington

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

QA1. How many epochs were required to train your perceptron on the 2-class Iris data having 2 features?

4 epochs

QA2. 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.)



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 don't believe there was any thrashing. The progress of the learning was positive from the chart as the separating lines and training data points match up pretty well and the model predicts the separation clearly.

QA3. What was the performance of your perceptron on the test data?

Only got 1 error, a percentage of 79 out of 80 items, giving the model a 98.75% success rate.

QA4. After plotting the ring data, describe its distribution in words.

The distribution is in a circular format with the blue and red dots around each other

QA5. Describe the sequence of separators obtained when training your perceptron for 5 epochs using the ring data. To what extent is there convergence? Thrashing? Hope for convergence?

There is no convergence as increasing the epoch number results in continuous weight changes and no convergence. There is thrashing due to change in slope and y-intercept and no hope for convergence

QA6. After you have re-mapped the ring data with the provided non-linear mapping function, plot the data and describe the distribution.

The distribution is separated between the red and blues as it is mapped in a way that allows good separation.

QA7. After training your perceptron on the re-mapped ring data, did it achieve convergence, and if so, how many epochs were used?

15 epochs were used to achieve convergence.

QA8. What does 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)?

Perceptrons seem to only classify data if a linear activation function. However, if it isn't used, sigmoid or ReLu in multiple layers, can allow the perceptron to classify data in non-linear ways.

QA9. Did you run into any difficulties either setting up for Part A or running the programs and answering the questions? If so, please describe them.

No.

QA10. What portion(s) of Part A did you find most worthwhile and why?

The most interesting and worthwhile part was seeing the workings of a perceptron, and how to classifies data and how it linearly separates out.