

Part A Report

Name: [Gordon McCulloh](#)

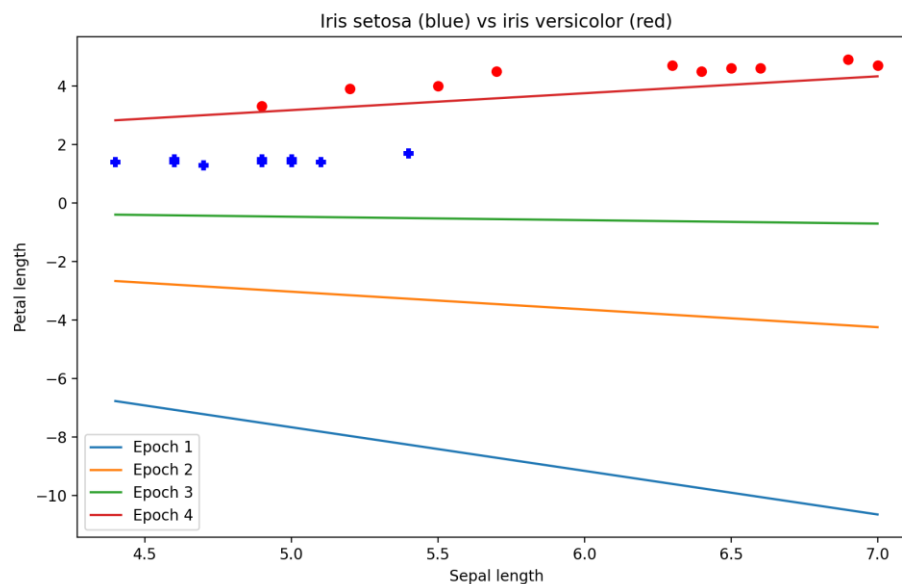
Assignment 6: Perceptron Classification and Training

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

A1. 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?

The training converged in 4 epochs. The perceptron had 2 errors on the test data out of 80 items. This represents a 2.5% failure rate for the 2-class Iris data.

A2. 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.)



A3. 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?

There is no thrashing as seen by the progress of the separator with each increase in epoch. The slope begins as negative but eventually finds a low positive value to correctly delineate the two data sets. The y-intercept also adjusts to meet the location of the data by translating upward from less than -6 to about 3. The agent learns by increasing the weight factor to correctly classify the setosa and versicolor.

A4. After plotting the ring data, describe its distribution in words.

The ring data appears to be evenly distributed in all four quadrants of the cartesian plane. There are 8 blue '+' marks, representing one data set, tightly grouped around the origin. There are also 24 red dots distributed in all four quadrants further away from the origin, representing the second data set. The two sets appear to form ellipses in cartesian coordinates.

A5. 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?

Training the perceptron for 25 epochs appears messy and unproductive. There is significant thrashing as the perceptron attempts to fit a linear separator to nonlinear data. The model never converges and will never converge unless the data is remapped because a linear separator cannot successfully delineate the given data in the cartesian plane.

A6. After you have re-mapped the ring data with the provided non-linear mapping function, plot the data and describe the distribution.

Remapping the data using the nonlinear mapping function produces a separable and evenly-grouped data set. The horizontal x-values range from -3 to 3, and the vertical y-values vary from about 1.5 to 3. The y-values represent radial measurements such that the first data set is grouped around 2 and the second above 2.5.

A7. After training your perceptron on the re-mapped ring data, did it achieve convergence, and if so, how many epochs were used?

Yes, the perceptron converged in 11 epochs. It gave 0 errors on the test data out of 32 items.

A8. 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)?

Perceptrons are unable to classify data that cannot be separated by a linear manifold. Clusters of data points of mixed sets represent a challenge because, if separable, they make it difficult to find a precise linear manifold. This is why the ring data required more epochs to converge than the iris data.