A6 Report

Name: Emma Cao

Assignment 6: Perceptron Classification and Training

CSE 415 Introduction to Artificial Intelligence, Autumn 2023, University of Washington

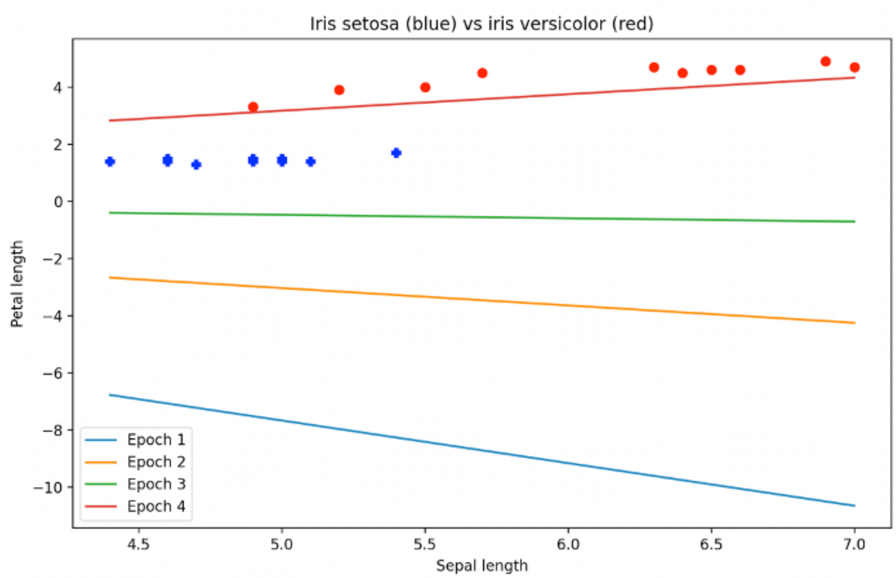
Please answer each question using text in Blue, so your answers stand out from the questions.

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. As the training progressed through these epochs, the decision boundary—illustrated by the lines—moved upwards, improving its ability to distinguish between the two classes. After four epochs, the decision boundary was able to linearly separate the data points effectively. However, there were a few instances where individual points deviated from the established trend, resulting in a couple of misclassifications within the test data points.

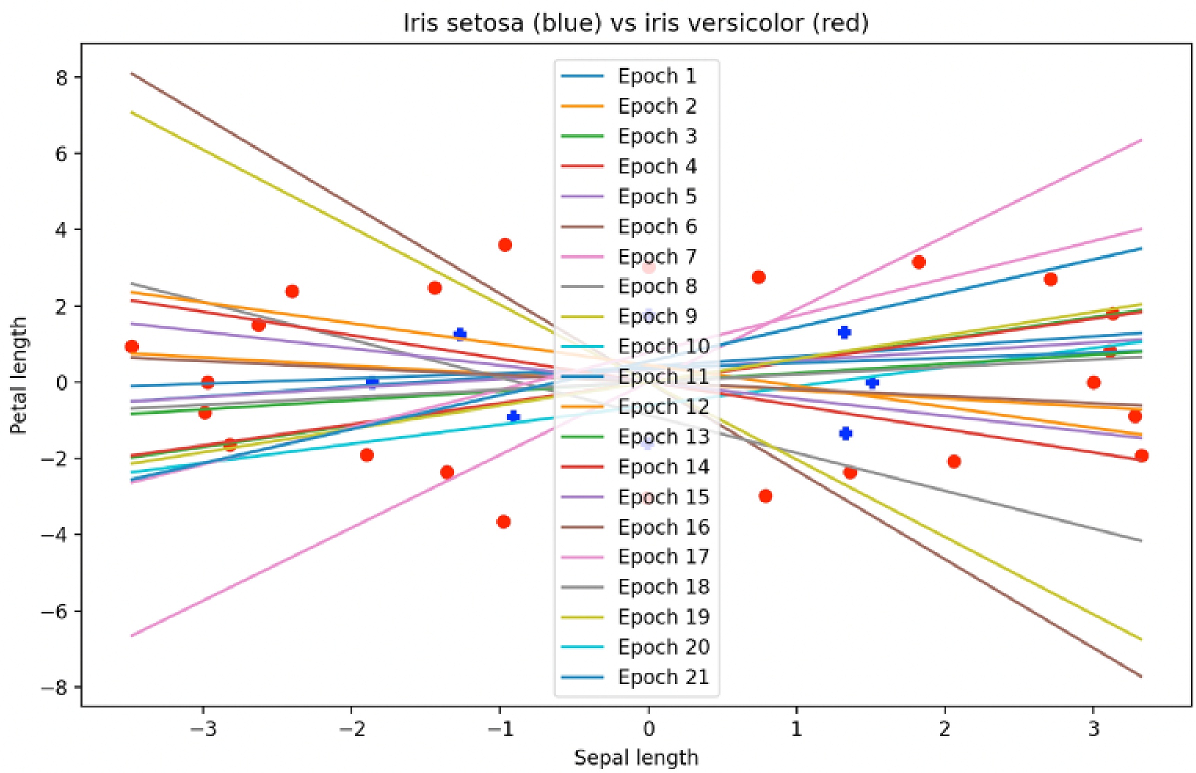
Q2. Include a graphic produced using matplotlib that shows both the training data points (in separate colors) and the "separating" lines implied by the weights at the end of each training epoch." (Reduce the graphic as necessary to make it fit here without taking up more than half the page.)



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 a methodical shift of the decision line with each epoch. The angle of the line consistently tilts upward, and the intercept gradually increases. This suggests a controlled and systematic training progression. There's no erratic movement indicative of instability in the learning process. The perceptron seems to be methodically fine-tuning its parameters, enhancing its classification capability between the two groups of Iris flowers. It hints at a well-calibrated learning rate, striking a balance between avoiding erratic jumps and ensuring the model's performance gets progressively better.

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

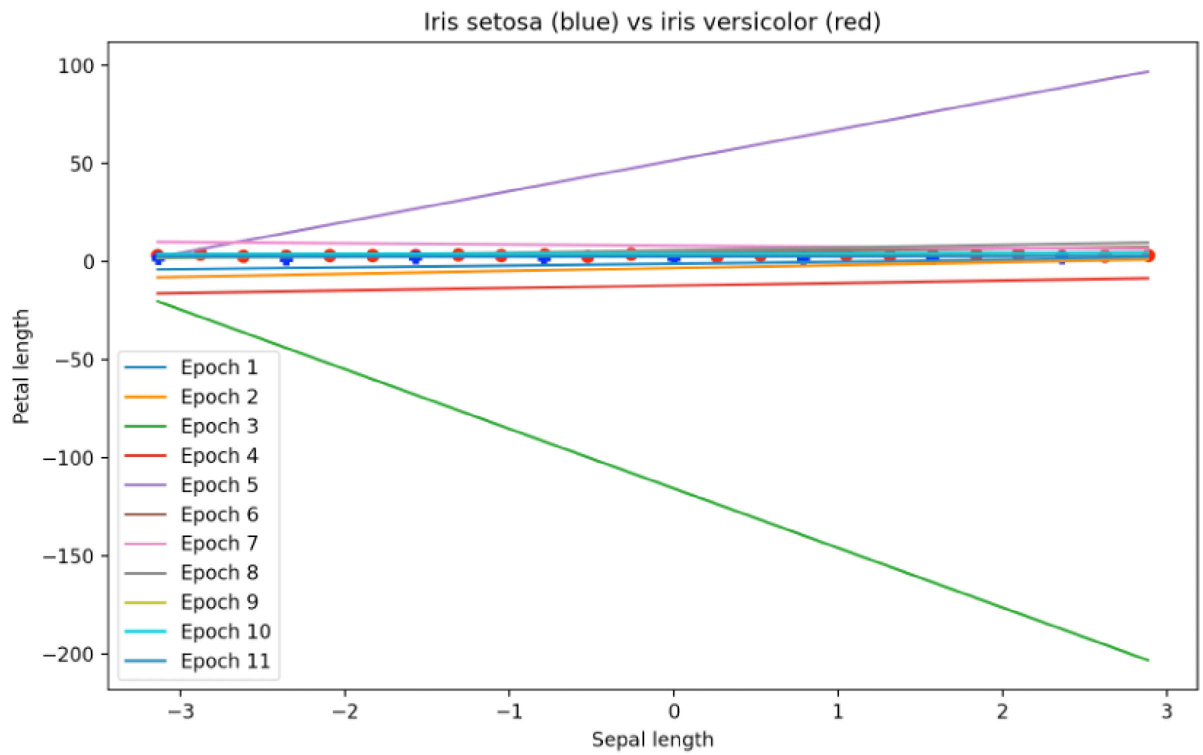


This diversity in the decision boundaries could indicate an unsettled learning process. The lines here cross over each other and spread across the plot, which is a sign of instability in the learning algorithm or thrashing. The distribution of the data points themselves seems to be reasonably separable by a linear classifier, as most of the setosa (blue) and versicolor (red) data points are clustered into two distinct groups, yet the perceptron has not stabilized on a consistent boundary that effectively separates the two classes throughout the epochs.

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?

The sequence of separators displayed on the plot, corresponding to the decision boundaries of a perceptron over 21 epochs, shows significant variability, with no clear sign of the model reaching a state of convergence. The lines representing each epoch's boundary are markedly different from one another, suggesting that the perceptron's weight adjustments are substantial and not yet stabilizing. This pattern is indicative of thrashing, where the perceptron fails to find a consistent solution, potentially due to a high learning rate or the non-linear nature of the data. Convergence is not evident in the provided epochs, as the decision boundaries continue to fluctuate without settling into a consistent divide between the two classes.

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



The plot illustrates the decision boundaries for a perceptron trained over 11 epochs. It does not show any further drastic changes, so the model has basically converged.

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

Yes, it achieved convergence, and 11 epochs were used.

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)?

If the data consists of clusters that cannot be separated by linear boundaries, the perceptron will have difficulty achieving convergence. This is because there is no straight line or hyperplane that can be drawn to separate different classes without errors. The perceptron will iterate endlessly, adjusting its weights at each epoch to reduce classification errors, but due to the nonlinear nature of the data, it will not identify a stable solution that correctly classifies all training samples. Therefore, sometimes tools like thermal mapper need to be applied to it to preprocess it.