Note: The stopping condition that we utilized for our learning problems compares the mean number of errors in our previous epochs and the mean number of errors in more recent epochs and stops training once the latter stops decreasing in comparison to the former.

T5.1

Experiment:

In task 2, we used our perceptron algorithm to learn a model for the three learning problems:

1. LP1: iris setosa (+1) versus not iris setosa (-1)
2. LP2: iris versicolor (+1) versus not iris versicolor (-1)
3. LP3: iris virginica (+1) versus not iris virginica (-1)

We presented our data in the same order as in iris.data and initialized all our weight vectors to 0.

Results:

1. In LP1, our model reaches zero errors after 4 epochs.
2. In LP2, our model never reaches zero errors, however,the number of errors stops decreasing at around 6 epochs at 3 errors.
3. In LP3, our model never reaches zero, however,the number of errors stops decreasing at around 4 epochs at 3 errors.

Analysis:

From these results, we can conclude that only LP1 is linearly separable since the number of errors reaches zero and the model converges.

T5.2

Experiment

In task 3, we once again use our perceptron learning algorithm to learn a model for the same three learning problems as task 2.

However, this time we perform this experiment three times under different conditions as specified below:

1. T3.1: Initializing weights to 1
2. T3.2: Initializing weights to four different numbers between 0 and 1
3. T3.3: Initializing weights with a set of random weights

For both experiments, we presented our data in the same order as the original iris.data file.

Results:

1. In T3.1, our results were as follows:
    a. In LP1, our model reaches zero errors in four epochs.
    b. In LP2, our model never reaches zero errors, however, the number of errors stops decreasing at around 6 epochs at 6 errors.

    c. In LP3, our model never reaches zero errors, however, the number of errors stops decreasing at around 4 epochs at 4 errors.
2. In T3.2, we set our initial weight vector as [0.1,0.2,0.3,0.4,0.5]. Our results were as follows:
    a. In LP1, our model never reaches zero errors, however, the number of errors stops decreasing at around 7 epochs at 2 errors.
    b. In LP2, our model never reaches zero errors, however, the number of errors stops decreasing at around 5 epochs at 5 errors.
    c. In LP3, our model never reaches zero errors, however, the number of errors stops decreasing at around 4 epochs at 3 errors.
3. In T3.3, our results were as follows:
    a. In LP1, our model reaches zero errors at 2 epochs.
    b. In LP2, our model never reaches zero errors, however, the number of errors stops decreasing at around 4 epochs at 6 errors.
    c. In LP3, our model never reaches zero errors, however, the number of errors stops decreasing at around 6 epochs at 4 errors.

Analysis:
From these results, we can conclude that our choice of initial weight vectors affects the training process. For example, the results of training of LP1 models across the three different experiments vary. In T3.2, LP1 does not reach zero errors before the mean number of errors stops decreasing while in T3.3, LP1 reaches zero errors in lesser epochs than it does in T3.1. Additionally, LP1 in T3.1 and T3.3 converge onto different sets of weight vectors.

T5.3

Experiment:
In task 4, we once again use our perceptron learning algorithm to learn a model for the same three learning problems as task 2.

However, this time we perform this experiment three times under different conditions as specified below:

1. T4.1: Shuffling data in iris.data file in a random manner
2. T4.2: Reshuffling data in iris.data file in a random manner

For both experiments, we set our initial weight to zero.

Results:

4. In T4.1, our results were as follows:
    a. In LP1, our model reaches zero errors in two epochs.
    b. In LP2, our model never reaches zero errors, however, the number of errors stops decreasing at around 36 epochs at 46 errors.
    c. In LP3, our model never reaches zero errors, however, the number of errors stops decreasing at around 47 epochs at 3 errors.
5. In T4.2, our results were as follows:
    a. In LP1, our model reaches zero errors in two epochs.

b. In LP2, our model never reaches zero errors, however, the number of errors stops decreasing at around 53 epochs at 39 errors.
c. In LP3, our model never reaches zero errors, however, the number of errors stops decreasing at around 38 epochs at 4 errors.

Analysis:
From these results, we can conclude that the order in which training instances are fed to the model affects the training process. For example, the results of LP2 across the two experiments greatly differed from their results in previous tasks. Additionally, LP1 in T4.1 and T3.2 converge onto different sets of weight vectors.