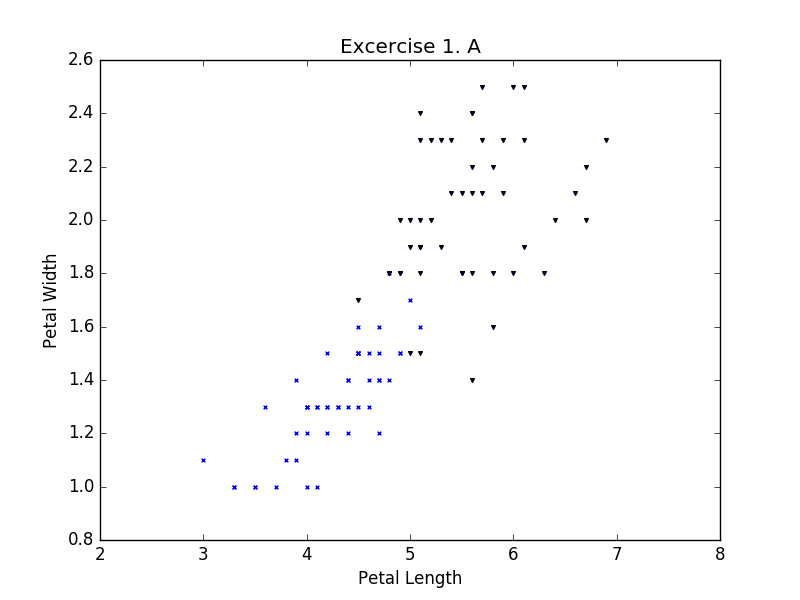
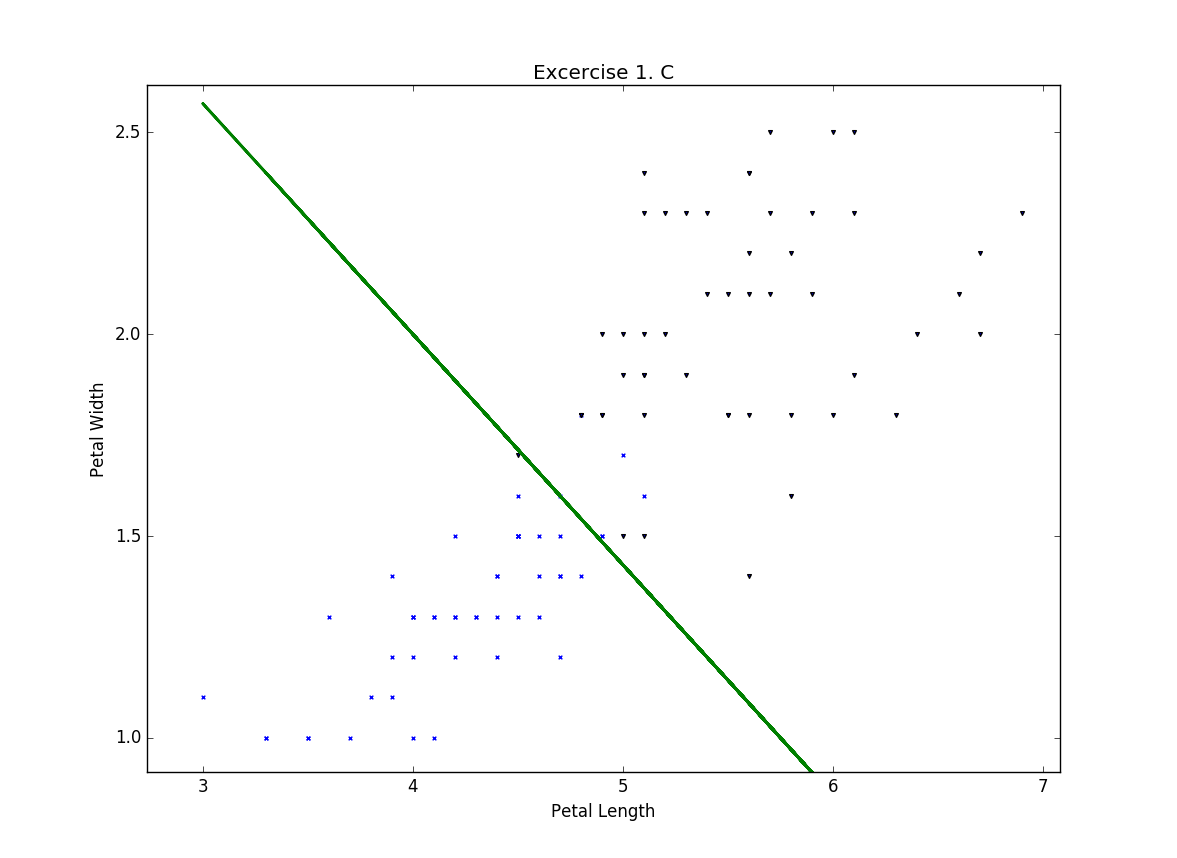
Huvra Mehta | HSM20 | Project 2

Exercise 1:

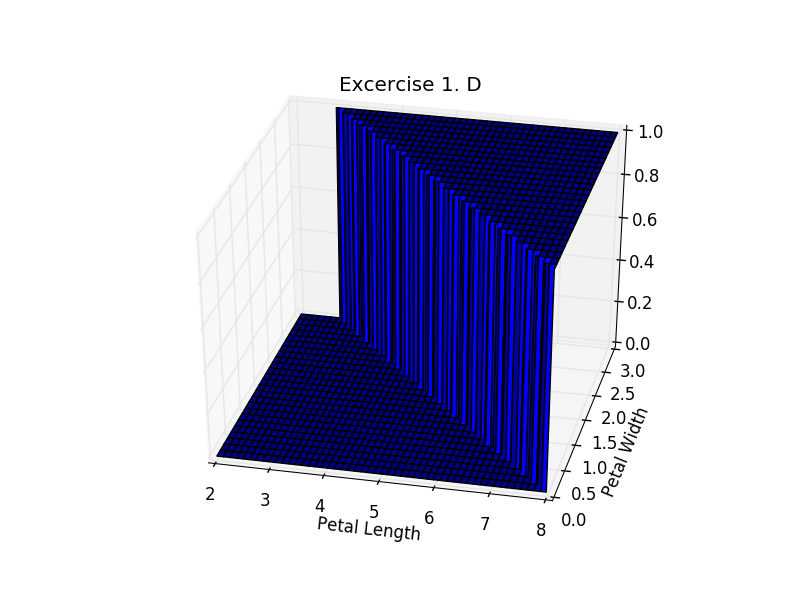
A) Plot 2nd and 3rd groups!



C) Plot dimension



D) 3D Plot



E)

Group 0: versicolor

Point: (3.3, 1.0)

Category: 0

Point: (3.3, 1.0)

Category: 0

Group 1: virginica

Point: (5.0, 1.9)

Category: 1

Point: (6.1, 2.5)

Category: 1

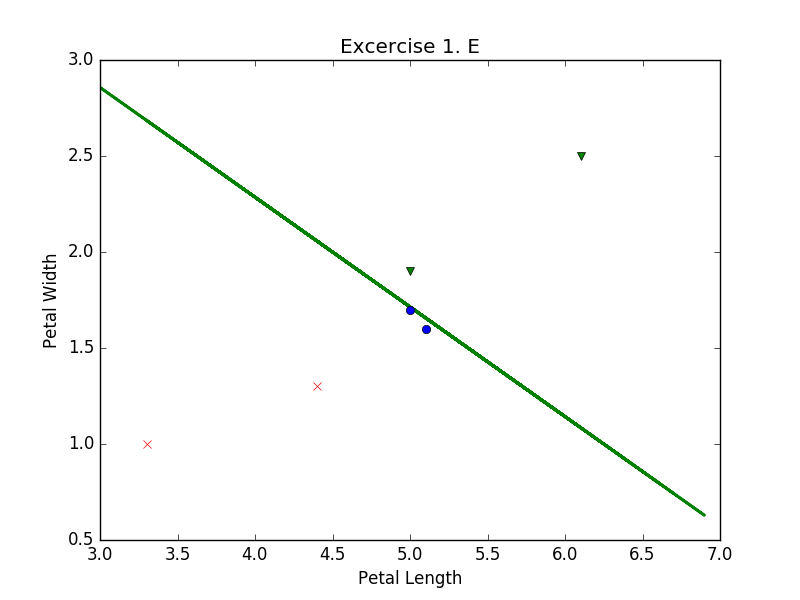
Near the Line

Point: (5.0, 1.7)

Category: 0

Point: (5.1, 1.6)

Category: 0



Exercise 2:

NOTE: For part A and B, I used a version of MSE that used the threshold stated in Question 1 where all values in the 2nd group are 0 and all in the 3rd are 1, allowing for neater numbers. But for part E, I used a version, which returned the pure sigmoid allowing for more gradation. It is this version that I will also use for question 3.

B)

Mean Squared Error for First Set of Weights:

W1 0.4

W2 0.7

BIAS -3.2

Mean Squared Error: 0.05 < Green Line

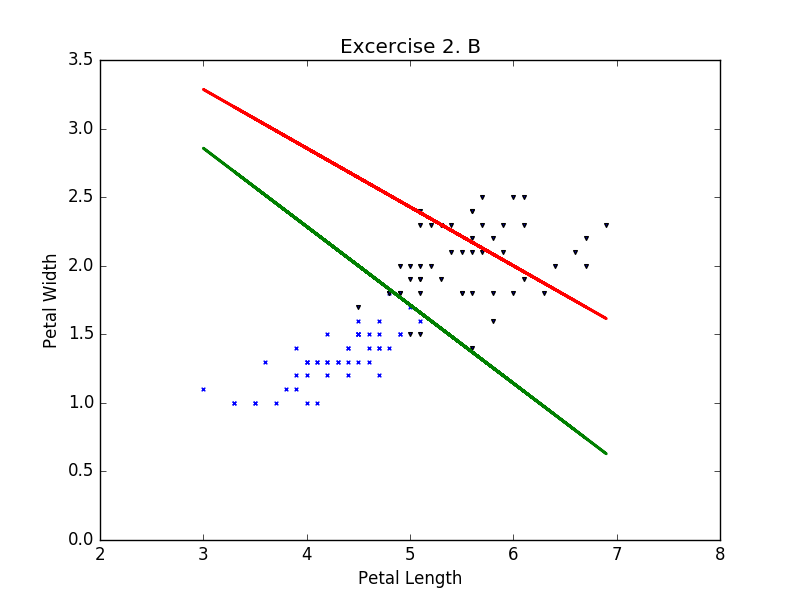
Mean Squared Error for Second Set of Weights:

W1 0.3

W2 0.7

BIAS -3.2

Mean Squared Error: 0.32 < Red Line



C) (On Paper)

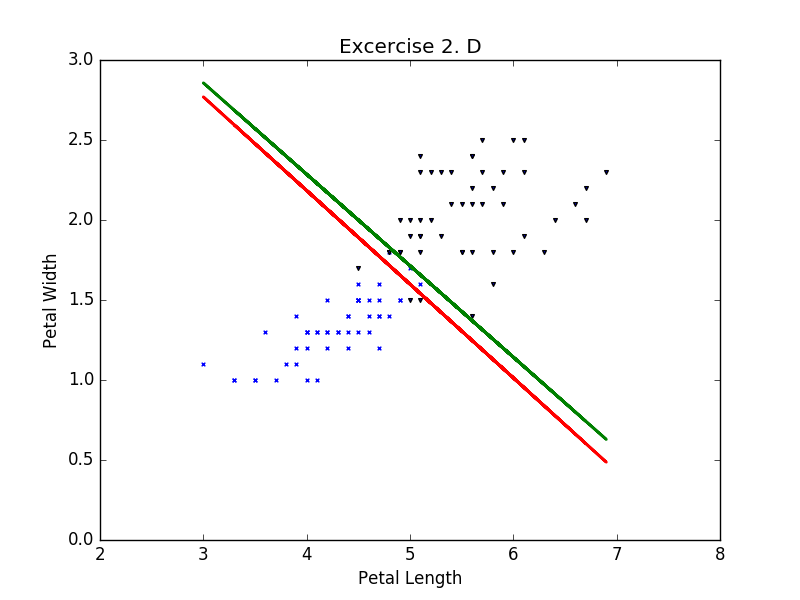
D) (On Paper)

E)

Mean Squared Error for First Set of Weights:

Old Weights: 0.4 0.7 -3.2

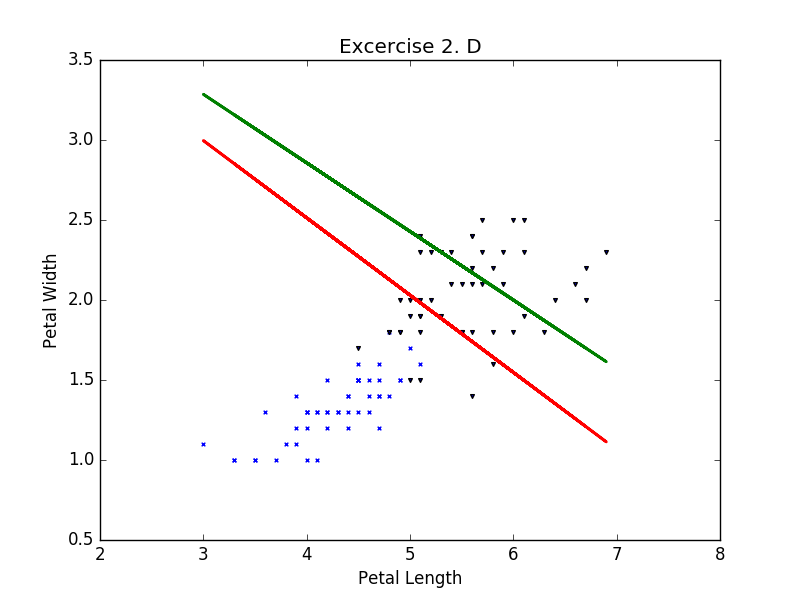
New Weights: 0.4138082888706862 0.7068078484415926 -3.2



Mean Squared Error for Second Set of Weights:

Old Weights: 0.3 0.7 -3.2

New Weights: 0.3468297708281769 0.7182007447547906 -3.2

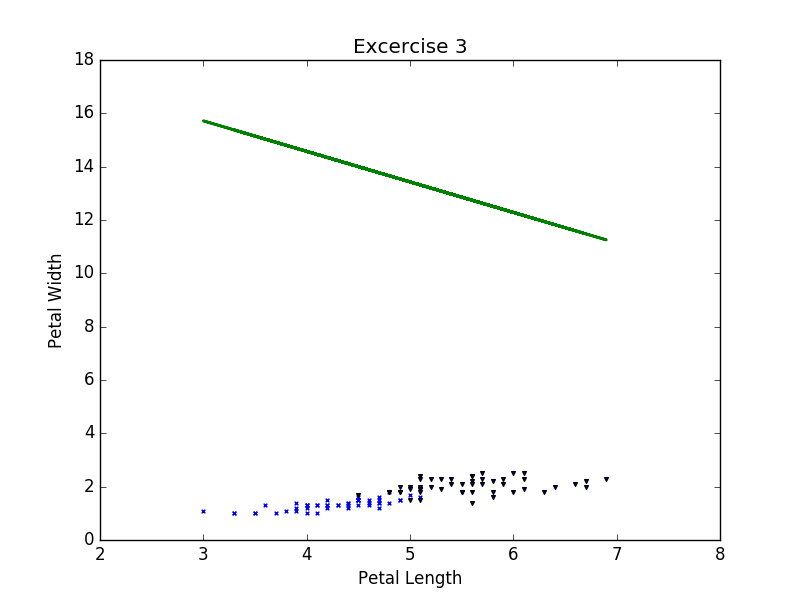


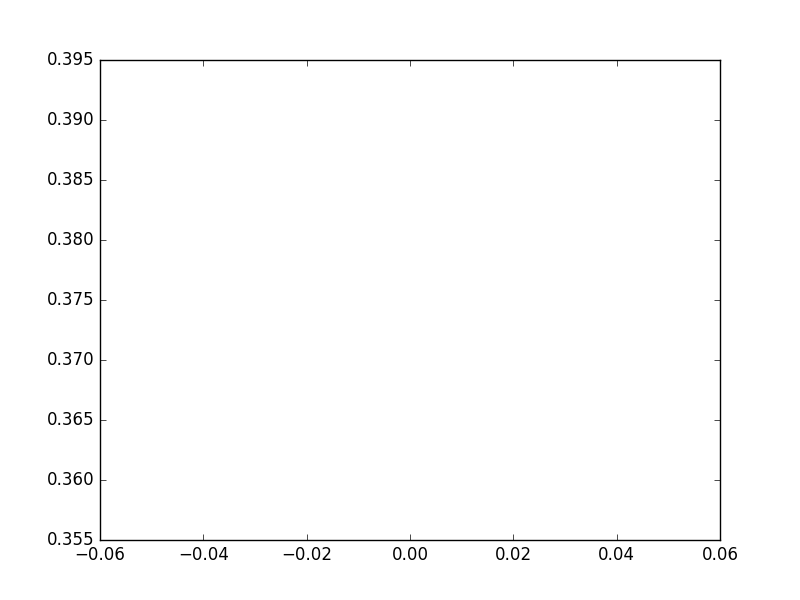
3)

Note: The dots are categorized based on their actual categories to show how the decision boundary is moving towards the correct answer that was defined in Problem 1 Part C.

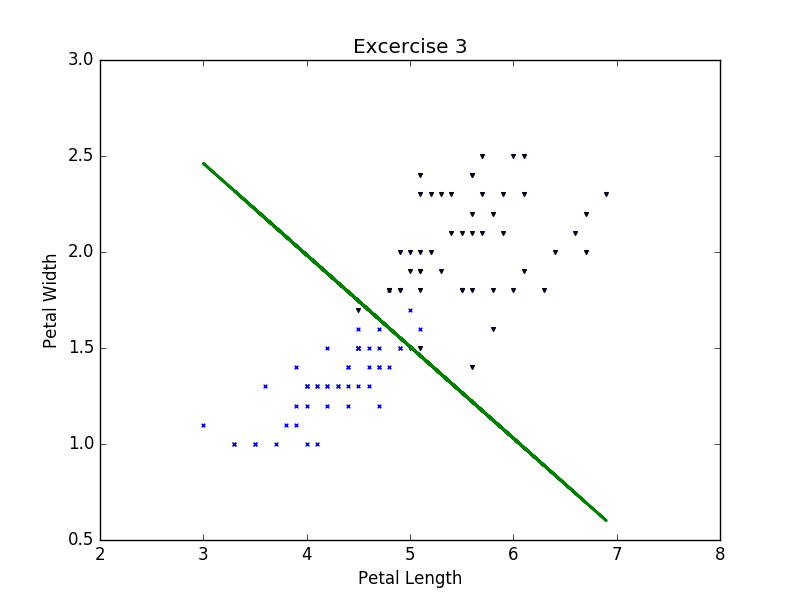
The learning curve plot keeps track of how the neural network gets better at categorizing the points over time.

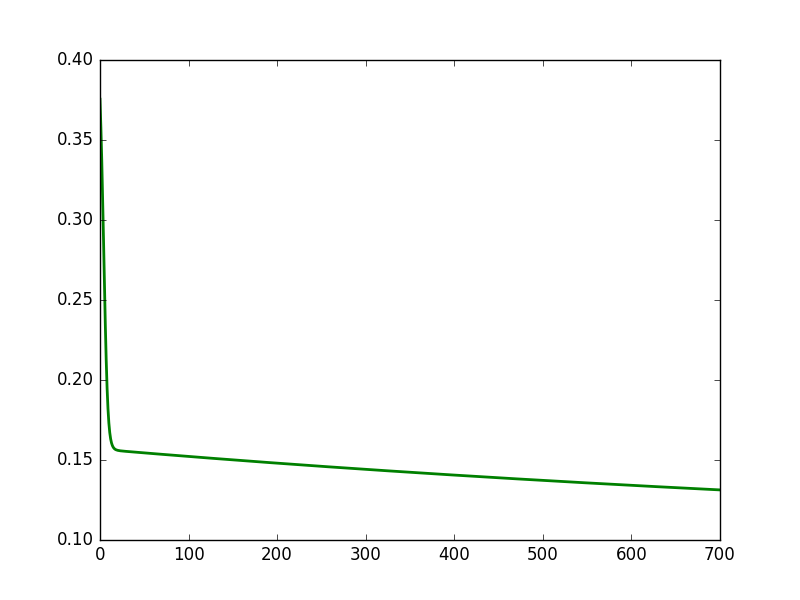
Initial:



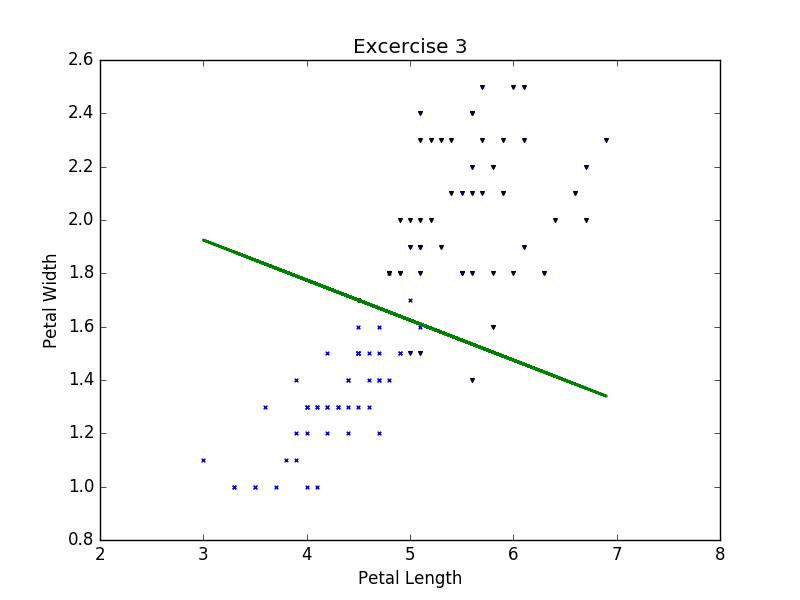


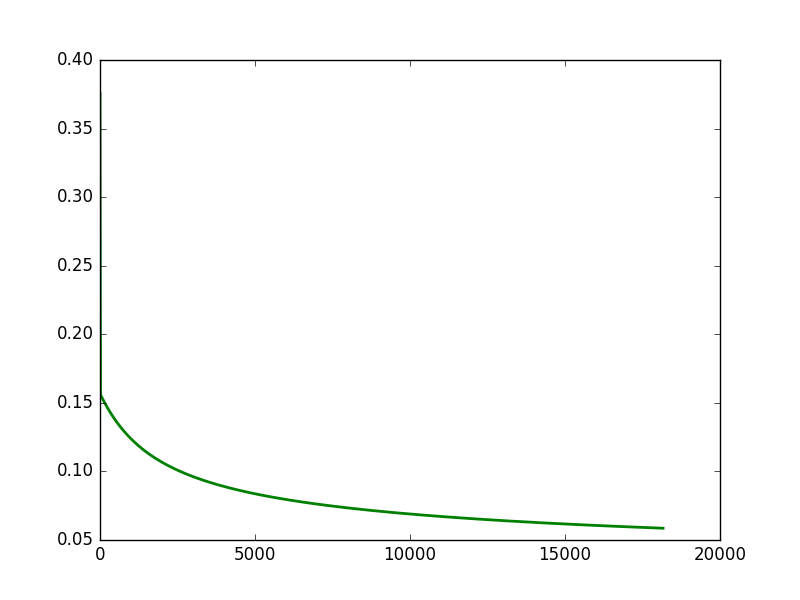
Middle:





Final:





D) Explain how you chose the gradient step size.

Using trial and error I started with .1 and decreased the step size until I was able to reach a step size that provided a smoother learning curve graph. I also wanted to ensure the run time would be manageable, around 10,000 tries. As such, I tried to keep my step size towards the larger side than the smaller.

E) Explain how you chose a stopping criterion

I stop my loop based on when the weights are no longer changing significantly with the assumption that once the weights had stopped shifting under a certain threshold then the algorithm had reached minimum.

To find this threshold I set my starting threshold incredibly small, around .00000000001, and ran the program, printing my weights and plotting my decision boundary for each iteration. After letting the program run for several minutes (as my incredibly small threshold could not be reached within five minutes), I paused it and took the last two weights that were printed and minus the new weight from the old weight for all three Ws (W1, W2, W0). I then took the smallest difference and set it as my new threshold. I think repeated this process again and again until my threshold could be reached within two minutes and I saw that the plot was no longer significantly changing after several iteration

As such, I decided upon the .0002 threshold.

Extra Credit!