Isabelle Chalhoub - u0678302
Karla Kraiss - u0830999
CS4300
December 8, 2017

A9 Lab Report

1. Introduction

In this lab we implemented perceptron learning and logistic regression to classify images of Wumpus World based on Pits, Wumpus, and Gold. We used sensor images to classify the weight vectors of features and predicted which class of image those vectors belonged to.

- How successful is perceptron learning compared to logistic regression given noisy data? (Success measured in predictions correct)

2. Method

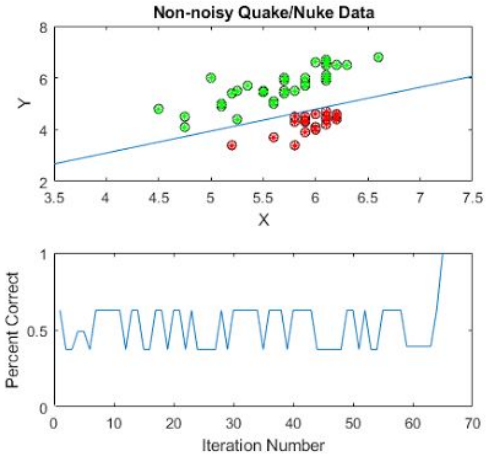
CS4300_perceptron_learning:
This method classifies images based on features and how much weight those features possess. Provided linearly separable data, perceptron learning will converge into a perfect solution using a weight update rule. It does this by predicting which class each image should go to and then checking the weights of the images against its prediction. If all its predictions were correct it exits, otherwise it continues this process until it has succeeded or reached max iterations.

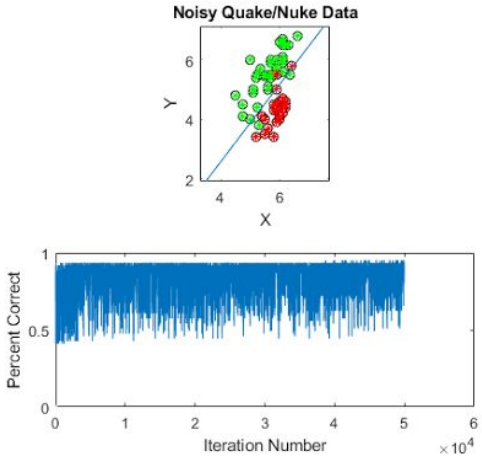
CS4300_logistic_learning:
This method also classifies images based on features and how much weight those features posses, but it does so with more interpretation. Logistic learning allows for more granularity near the linear classifier boundary. It does this by minimizing the loss of weight data by keeping track of the prediction error, as well as the prediction correct. We check predictions by measuring the error. If there was no error, we exit.

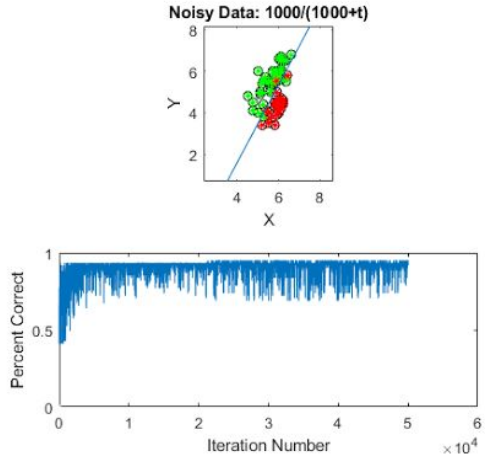
3. Verification of Program

For verification we compared our formulas output with the quake data to the book's output. We were only able to compare the prediction graphs, but those graphs showed our correctness.

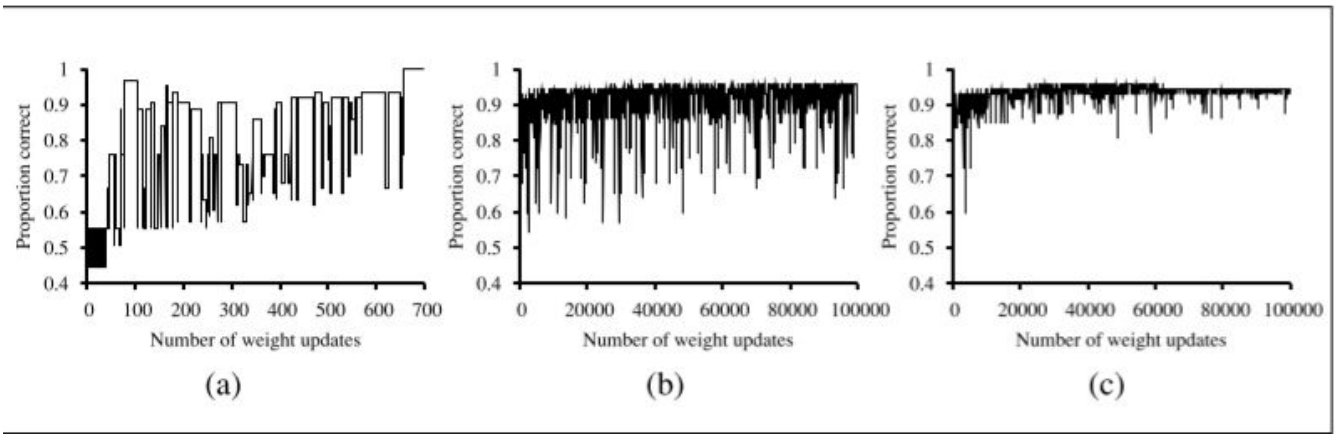
Our Perceptron Learning:



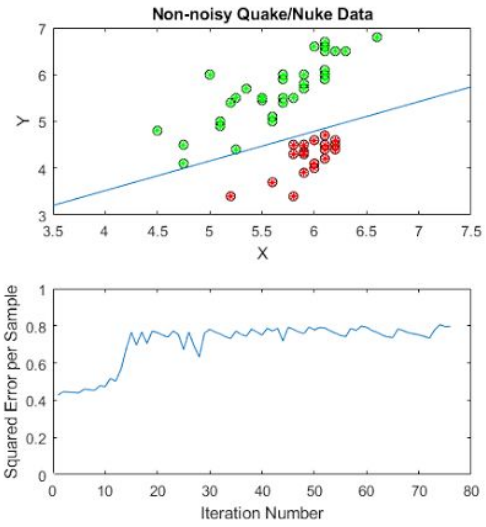


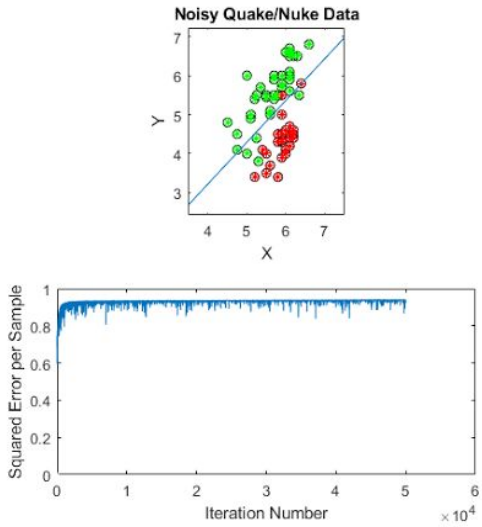


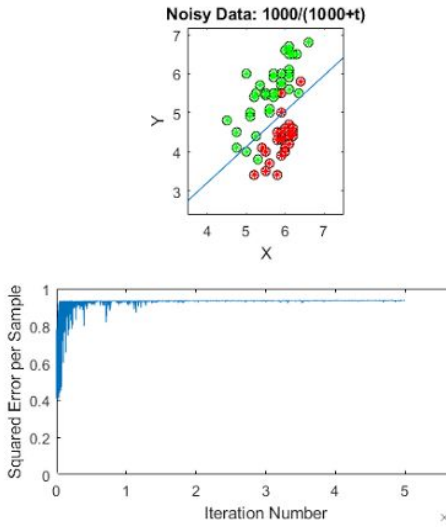
Books Perceptron Learning:



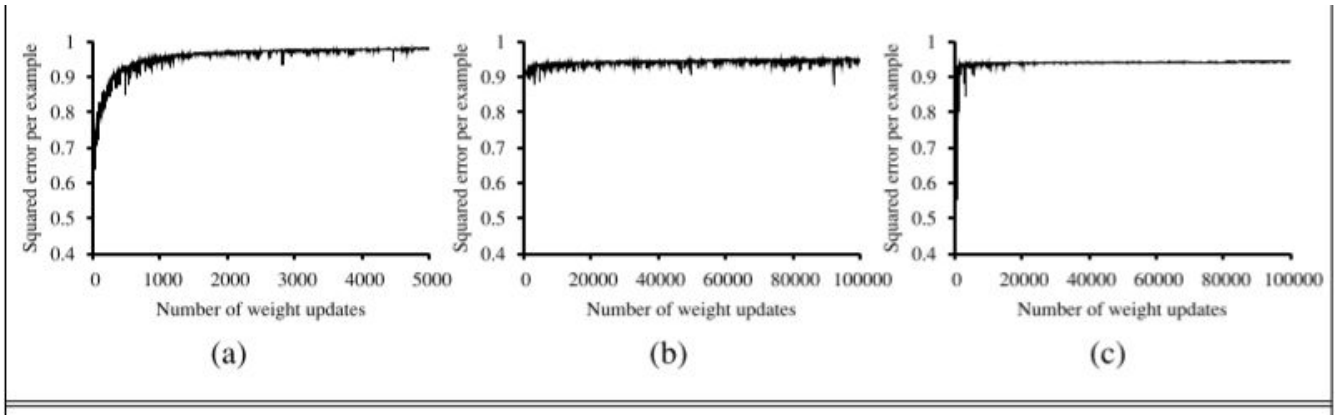
Our Logistic Regression:



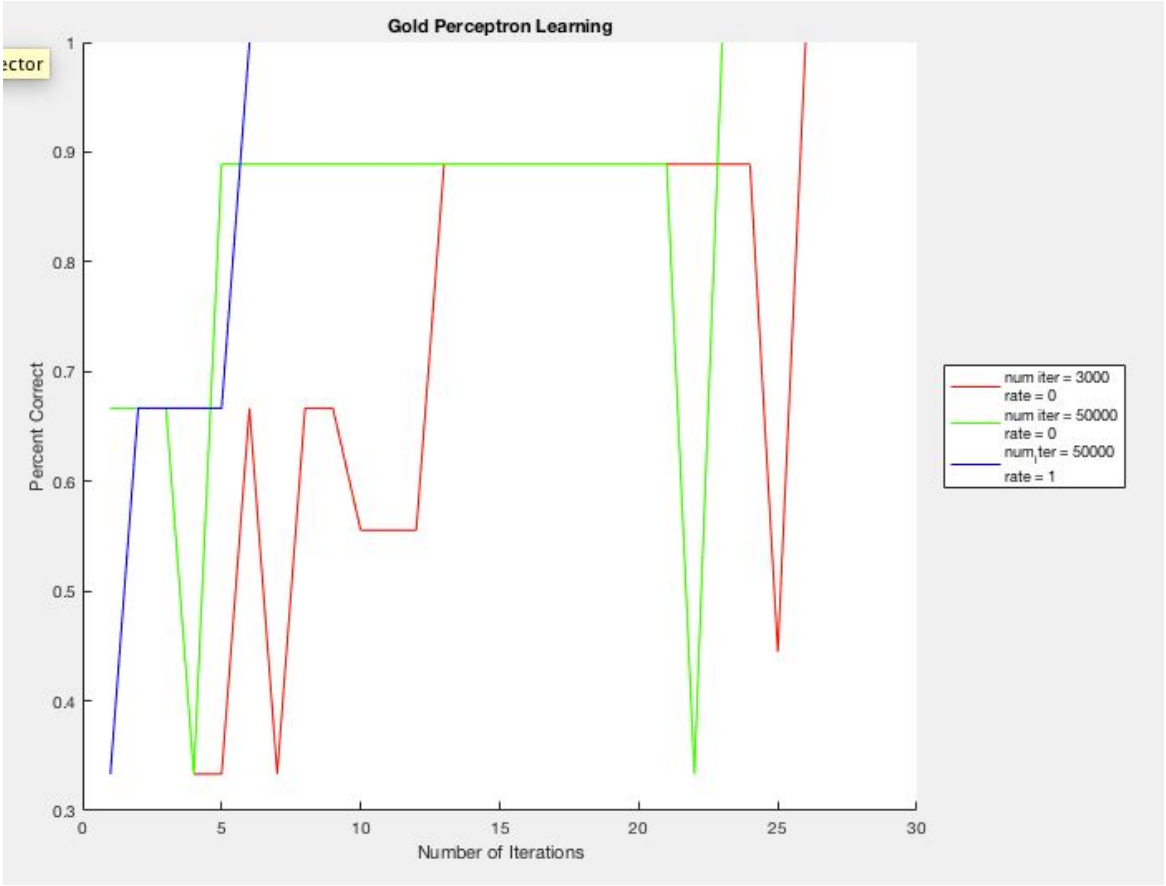


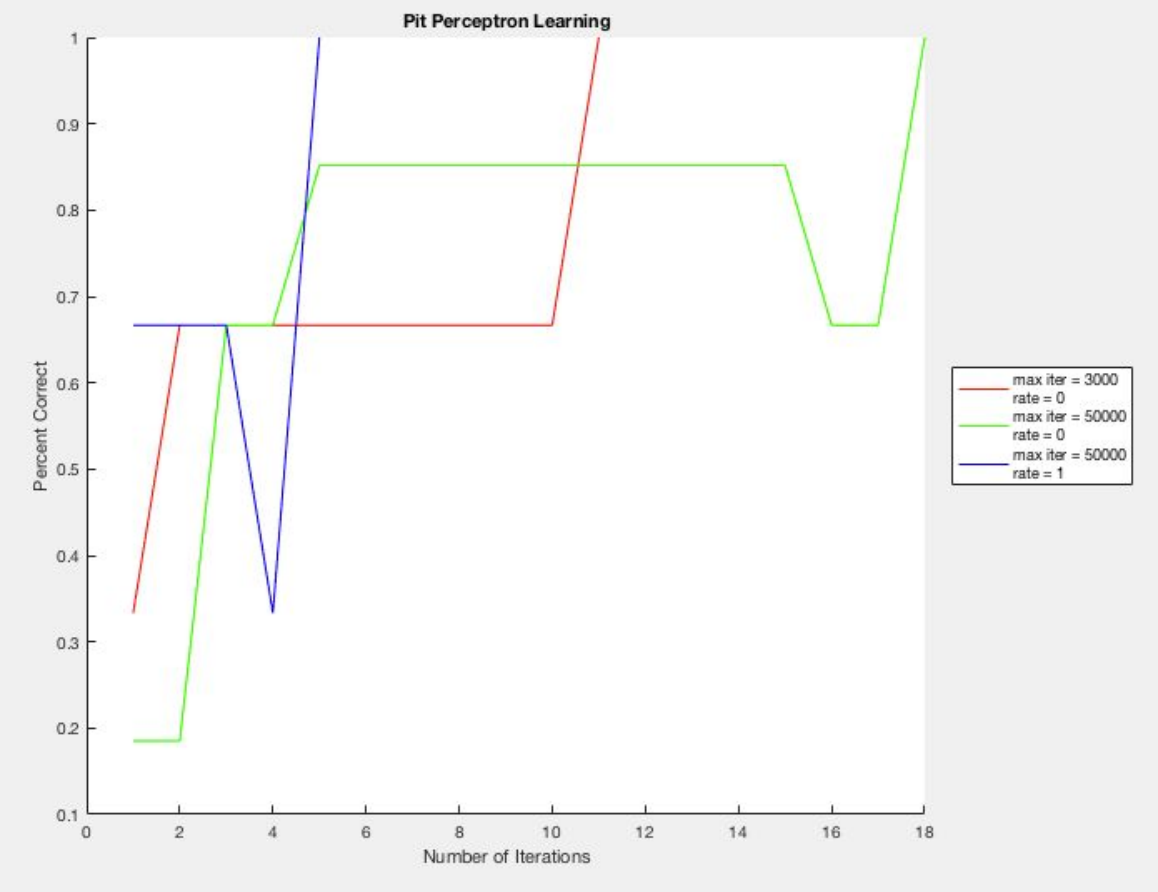


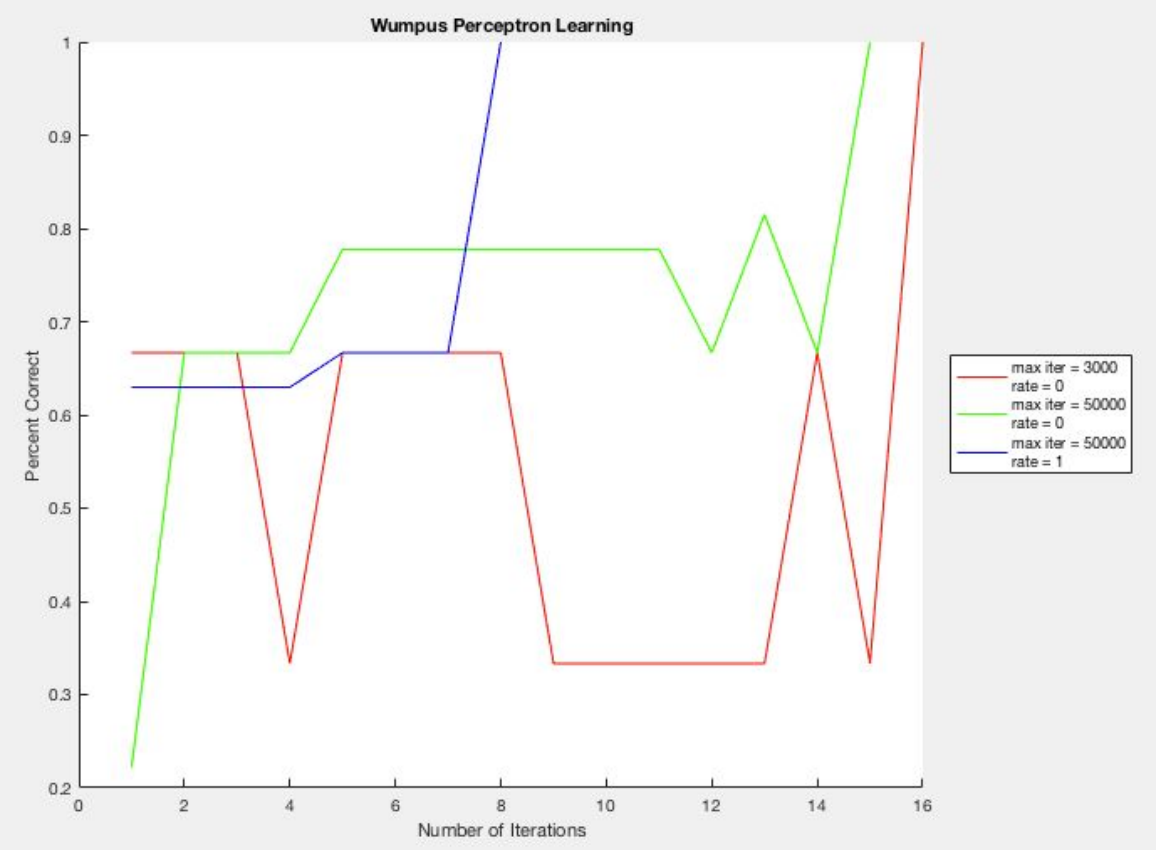
Book’s Logistic Regression:

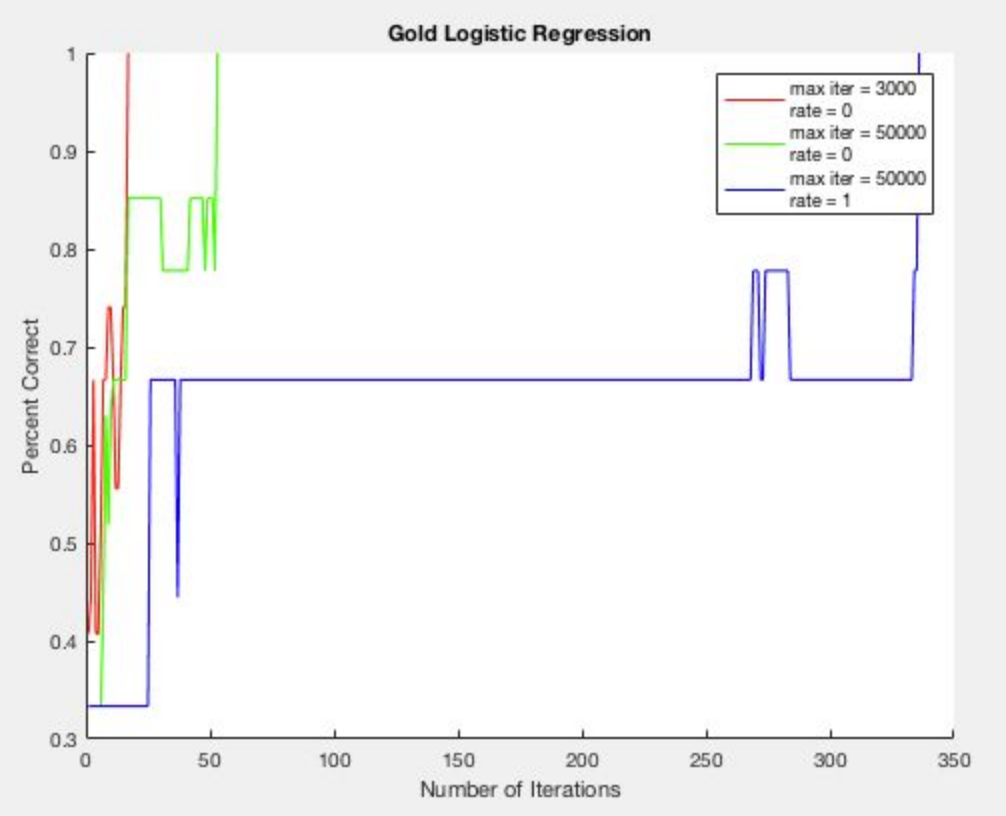


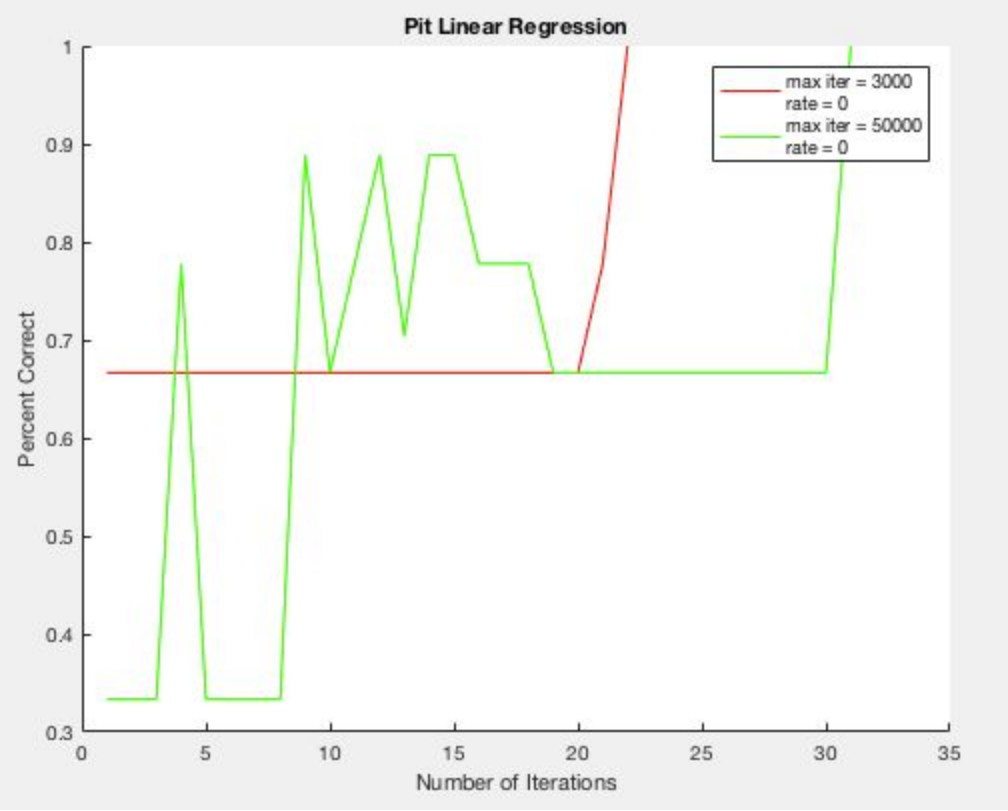
4. Data and Analysis

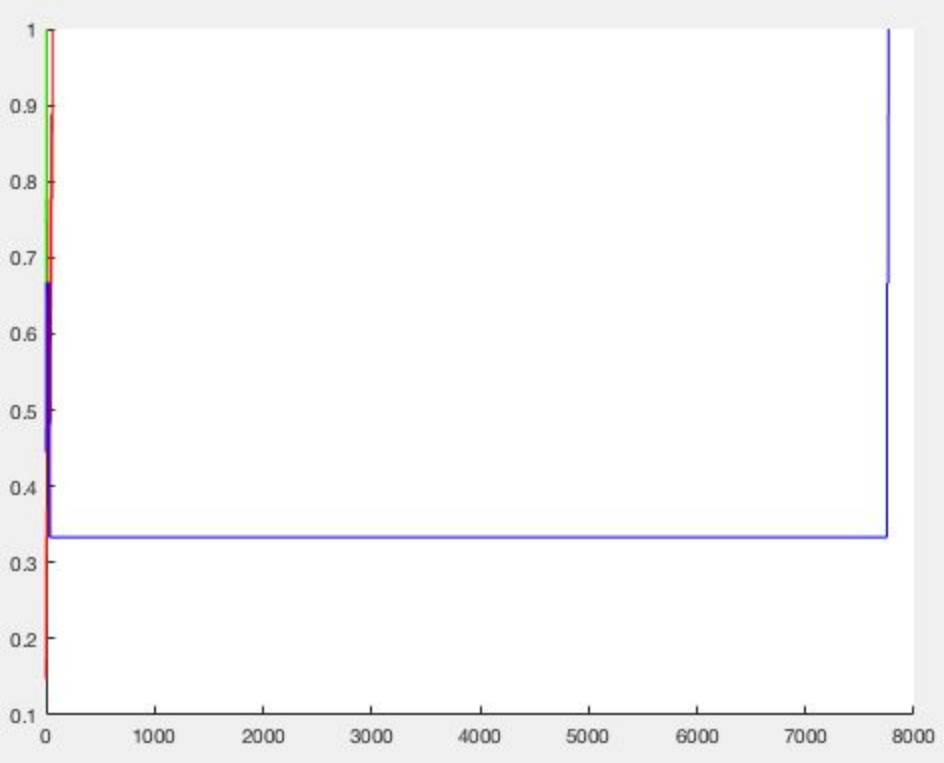



Pit Perceptron Learning
Percent Correct
Number of Iterations
max iter = 3000
rate = 0
max iter = 50000
rate = 0
max iter = 50000
rate = 1

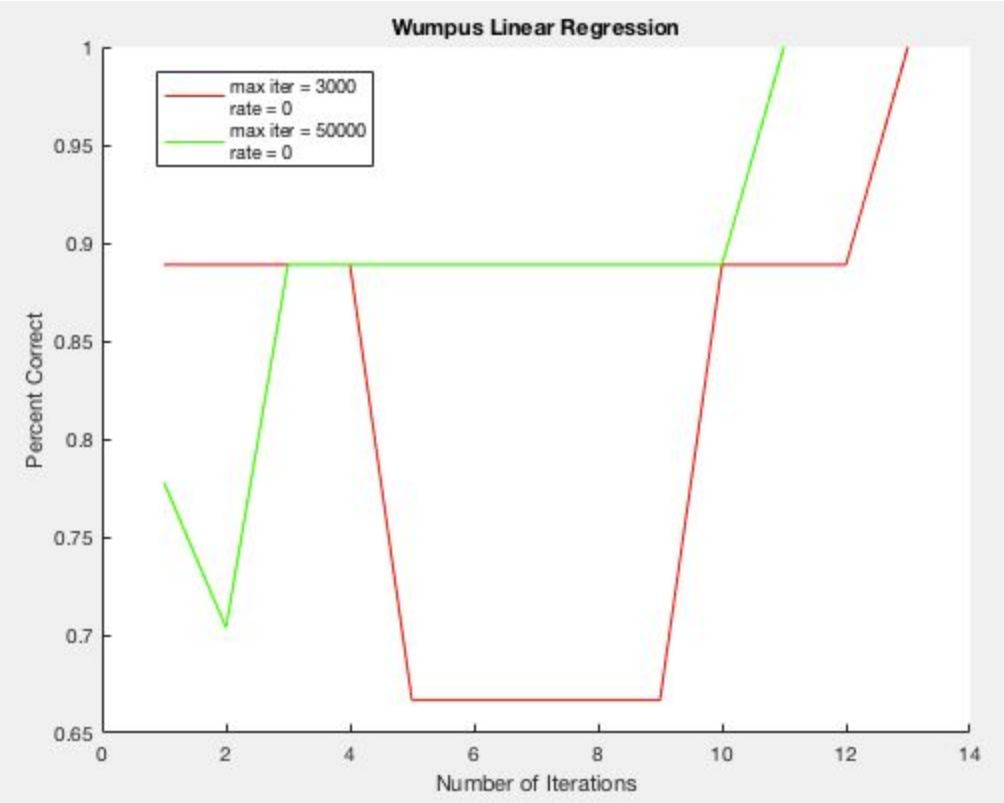

Wumpus Perceptron Learning
Percent Correct
Number of Iterations
max iter = 3000
rate = 0
max iter = 50000
rate = 0
max iter = 50000
rate = 1

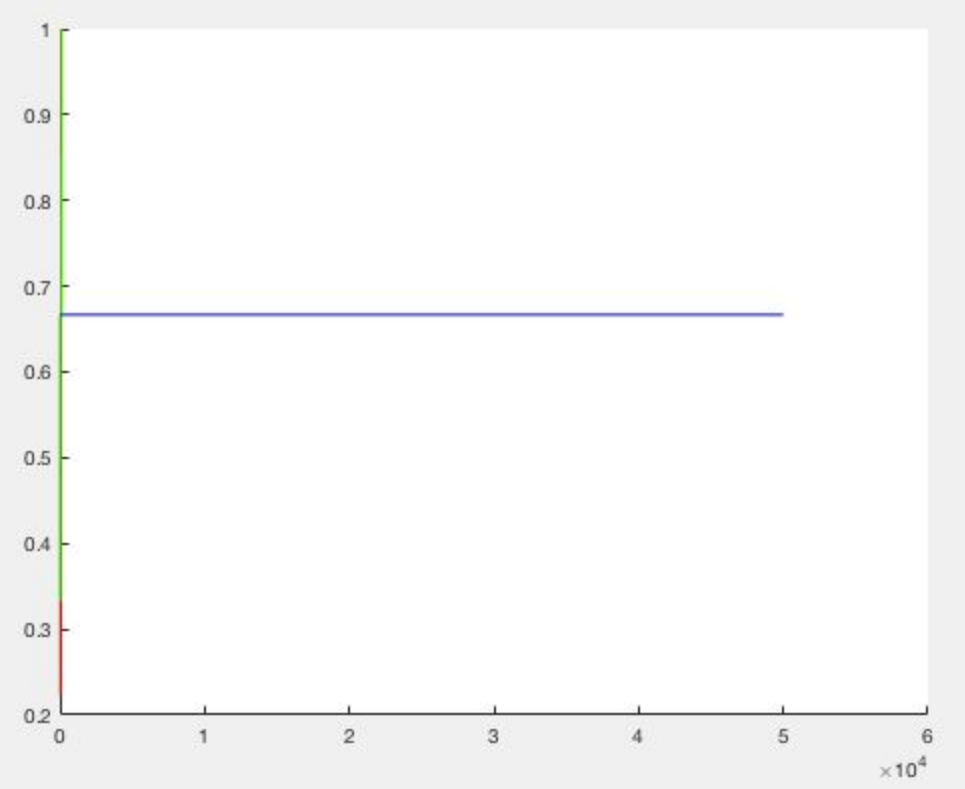

Gold Logistic Regression
max iter = 3000
rate = 0
max iter = 50000
rate = 0
max iter = 50000
rate = 1
Percent Correct
Number of Iterations
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0
50
100
150
200
250
300
350


Pit Linear Regression
max iter = 3000
rate = 0
max iter = 50000
rate = 0
Percent Correct
Number of Iterations
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0
5
10
15
20
25
30
35



The blue line in this graph is max iter = 50000 and rate = 1 for Pit linear regression. It had to be graphed separately because it takes so long to converge to 1.





Similarly, this is the Wumpus Logistic regression graph specifically to show that when the max iter is 50000 and rate is 1, it rarely converges to 1 and often stops due to hitting its max iteration.

5. Interpretation

Perceptron learning does a better job of predicting image classes when the features are more distinct. Because it does not also calculate the error and take it into consideration on the next iteration, it is able to converge more quickly and definitely. The problem is that perceptron learning is either 0/1 for binary classification, and does not allow for any granularity (such as 0.5).

Logistic regression does a much better job of predicting image classes given “noisy” data. Meaning when the features do not have a large weight or distinctness to them, it will still put them in the correct class. This is because logistic regression allows for more granularity near the boundary of of the linear hyperplane.

Both these characteristics can be found in our data sets, especially as iterations are increased. What was especially interesting was when rate was enabled so that the alpha was incorporated logistic regression took many more iterations to converge, if ever. This is

most likely because Logistic regression allows for more granularity, making it more difficult to fully converge given finer points to converge to.

6. Critique

In this assignment we had to figure out a rudimentary approach for image classification using a binary classifier. An important concept we learned was to keep each part of the formula compartmentalized. It was very easy to get overwhelmed when looking at it all at once, but when each small section was considered and analyzed we were able to build up the formula correctly. This also allowed us to better understand the differences between the two approaches.

Something we could improve on if doing this assignment again would be using more diverse data sets. The noisier a dataset was, the more it showcased the difference between logistic regression and perceptron learning.

7. Log

Isabelle's Log: (Odd Sections)
About 7 hours.

Karla's Log: (Even Sections)
I spent about 10 hours on this assignment. 7 hours coding, and 3 hours on the lab report.