A9 – CS4300 Assignment A9 Lab Report

Artificial Neural Networks

By Haochen Zhang & Tim Wei
12/10/2017

1. Introduction

In this assignment, we were to compare perceptron learning and logistic regression on image classification. Two method were implemented; they are, respectively, *CS4300_perceptron_learning* and *CS4300_logistic_learning*. In this report, we want to answer the following points of interests:

- What is the difference between perceptron learning and logistic learning on the aspect of learning pattern?
- Comparing rate is 1 and alpha is changing every step to rate is 0 and alpha is an constant, how does the performances differ?

2. Method

There are two parts in this lab which are image processing part and machine learning algorithm.

For image processing part, we use *imread* to read a jpg file into an matrix. Then, we pass our matrix into a function called *imresize* to reform it into a 15 by 15 matrix. After that, we transform each cell from values 0~255 into binary format. We parsed a total of 27 images (9 w, 9 g, 9 p) and store it in an array as our training example. We also create a label space array with 27 rows containing 27 labels representing as 1 for g, 2 for p and 3 for w.

For learning algorithm part, we implement perceptron learning algorithm and logistic learning algorithm. In *CS4300_perceptron_learning*, we are using

$$w = w + ((alpha * (y\_ - h)) * x\_)'$$

as its loss function where $h = (x\_ * w) >= 0$.

The biggesting difference between *CS4300_logistic_learning* is its loss function which is replaced by

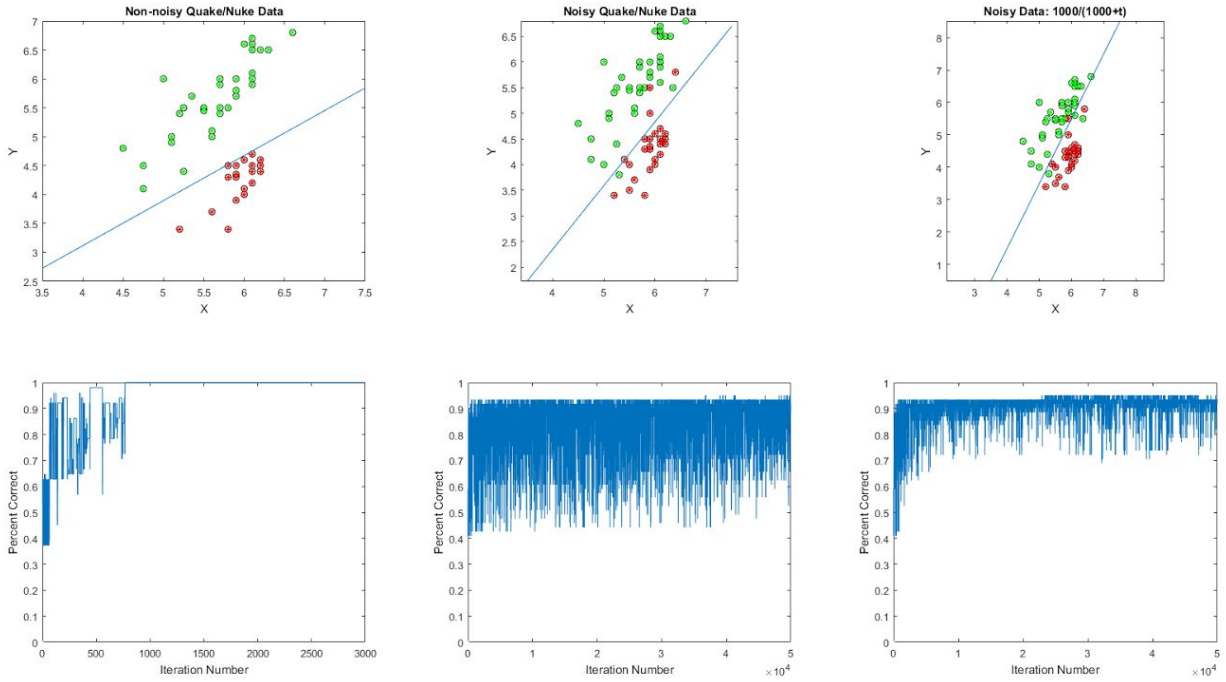
$$w = w + (alpha * (y\_ - h) * h * (1 - h)) * x\_'$$

where $h = 1/(1 + \exp(-x\_ * w))$.

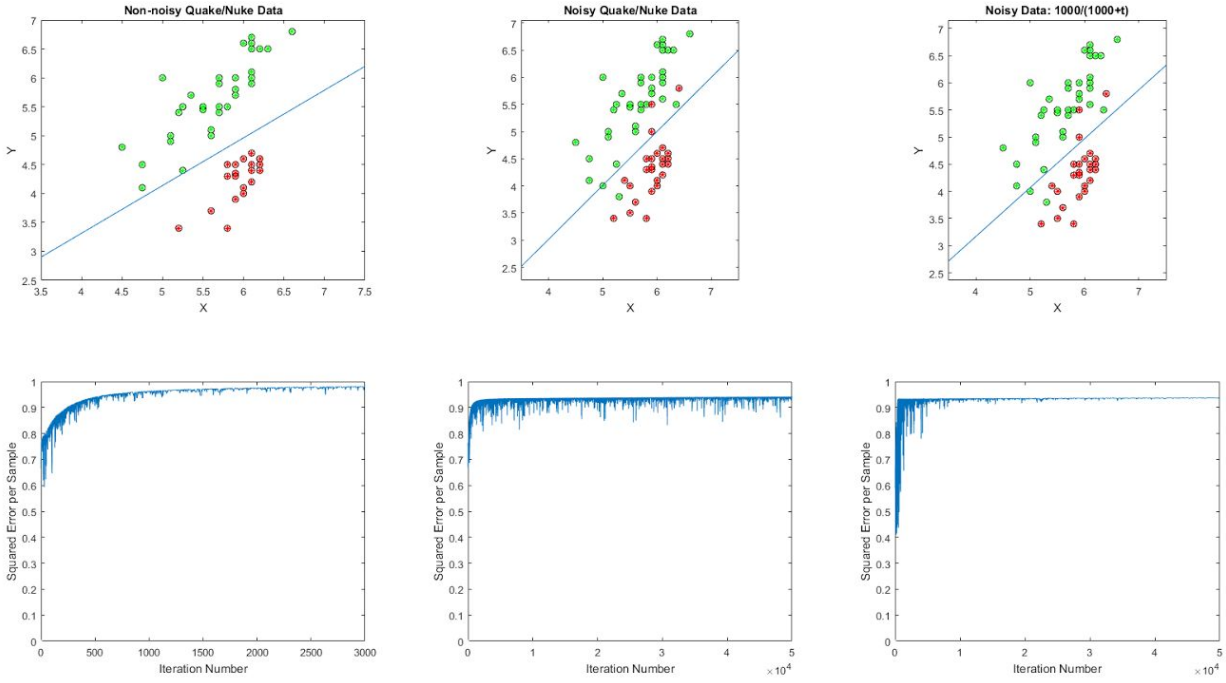
While training our data, we pass our label space with a logical equal. For example, if we are trying to classify image with letter “g”, our second parameter would be *y == 1* instead of just y. Thus, we will have a binary format matrix as our label space.

3. Verification of Program

We used *CS4300_demo_perceptron* and *CS4300_demo_logistic* (replacing *CS4300_PL* with *CS4300_perceptron_learning* and *CS4300_LL* with *CS4300_logistic_learning*) to create the following plots. These plots resemble to Figure 18.15 R & N p. 723 and Figure 18.16 from p. 725. Also the classifier linearly separated the data desirably.



Perceptron learning algorithm



Logistic learning algorithm

4. Data and Analysis

We ran *CS4300_perceptron_learning* and *CS4300_logistic_learning* to classify G, P, and W 1000 times each. The table below shows the result.

		CS4300_perceptron_learning			*CS4300_logistic_learning*		
		G	P	W	G	P	W
rate = 0	mean	20.7760	5.2470	11.7350	212.2000	81.3180	46.3330
	variance	152.5644	18.6346	67.5463	2484.2	633.2701	379.8920
	95% CI	20.7760 ±0.7656	5.2470 ±0.2676	11.7350 ±0.5094	212.2000 ±3.0892	81.3180 ±1.5597	46.3330 ±1.2081
rate = 1	mean	19.8750	5.6410	11.4230	47.3580	19.8340	13.5210
	variance	134.1675	19.3355	59.3054	180.4903	51.3818	28.6082
	95% CI	19.8750 ±0.7179	5.6410 ±0.2725	11.4230 ±0.4773	47.3580 ±0.8327	19.8340 ±0.4443	13.5210 ±0.3315

5. Interpretation

- What is the difference between perceptron learning and logistic learning on the aspect of learning pattern?

Compare to *CS4300_perceptron_learning*, *CS4300_logistic_learning* tend to converge more smoothly. But *CS4300_logistic_learning* takes more iterations to do so.

- Comparing rate is 1 and alpha is changing every step to rate is 0 and alpha is an constant, how does the performances differ?

Changing rate shows no significant difference for *CS4300_perceptron_learning*, but for *CS4300_logistic_learning* there is a large fall in mean iteration taken and its variance.

6. Critique

We are only using 27 example as our training dataset in this lab. It way too small to be an effective instance space. Furthermore, we don't have any test data to verify correctness and check overfit. To move further on this lab, we should use k-fold cross validation to determine the best hyper-parameter.

For the image processing part, we only set the middle column as our features. To proceed, we could the whole image or even analyze colorful image.

7. Log

Haochen Zhang(Section 2, 4, 6)
A total of 5 hours was spent performing the experiment in Matlab.
A total of 3 hours was spent performing the experiment in writing the report.

Tim Wei (Section 1, 3, 5)
A total of 5 hours was spent performing the experiment in Matlab.
A total of 3 hours was spent performing the experiment in writing the report.