CSE 353
HomeWork 3

JeongYoon Lee (114133199)

a) **Introduction. Brief summary of what you think the assignment is about**

This assignment is about implementing the perceptron learning algorithm(PLA) and pocket algorithm to classify data linearly.
There's various data with linear separable data, and non linear separable data. I'm looking forward to finding proper classifier with this algorithm, and for the non-linear separable data, I'll find the best model with smallest error after enough iteration.

After enough iteration, I'm hoping to find the error rate will converge at some point, and that will be the optimal value.

b) **Method. Brief outline of your (algorithmic) approach**

Perceptron learning algorithm(PLA),

Initialized W = [0,0]
for iteration n,
     if(sign(wT * Xn) != Yn): # This is mistake!
        W(t+1) = W(t) + Yn*Xn # update weight value

     If there's no error anymore, return W(t+1)

Training error will be ((the number of points that classified wrong) / (the number of whole data)).

Pocket algorithm is similar with PLA, but the only difference between them is there's no exact answer for weight for this nonlinear separated data. Since most data of this homework cannot divided into two parts with linear line, we should find the best weight value which minimize error rate. I will also try to print each plot whenever the weight updated, so we can track the learning process through this. You can check this process in my python file.

I also try to find optimal iteration number through printing error rate for each iteration in dataset 2 model.

c) **Experiments. Tables and/or pictures of intermediate and final results that convince us that the program does what you think it does.**

1. PLA with Linear Separable data
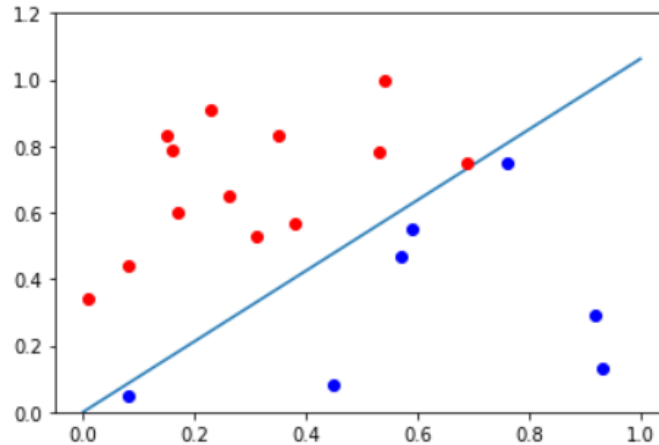
(a) Algorithm : PLA

Dataset : X (3*20), Y(1*20)

Iteration : Until find best weight value which classifing data well

Training error rate : 0%

Optimal weight vector = `[2.199999999999997 -2.069999999999994]`

(b) Some iteration and final classified data

```
Training error rate :   0.0
success
weight vector:  2.1999999999999997 -2.0699999999999994
```


[Final classified data plot]

(c) Training error rate : 0%

2. Pocket PLA with Non Linear Separable data

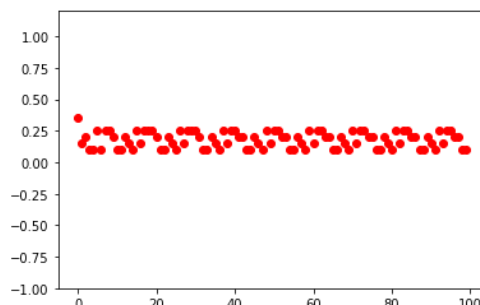(a) Algorithm : Pocket

Dataset : X (3*20), Y(1*20)

Iteration : Until find optimal weight value which classifing data well

Training error rate : 0.1

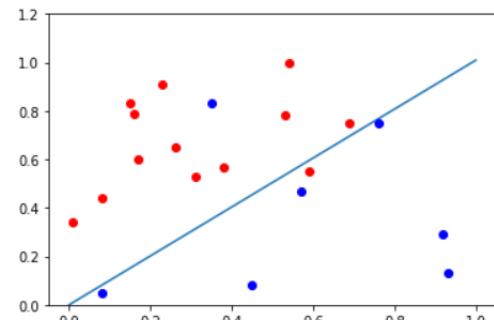Optimal weight vector = `[1.0699999999999998, -1.06]`

I tried to find optimal iteration number through printing graph of error rate for each iteration, but the result is not that meaningful. The error rate keep up and down while doing iterations. So I decided to choose 100 iterations for this model.

(b)

[Error rate based on iteration]   [Final classified data result]

(c) Training error rate : 0.1

3. Pocket PLA with Linear Hand Crafted data

(a) Algorithm : Pocket

Dataset : X (3*1561), Y(1*1561)

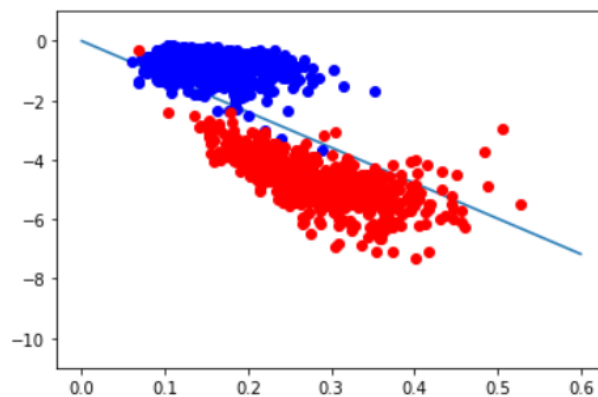Iteration : Until find optimal weight value which classifing data well

Training error rate : 0.0294

Optimal weight vector = [29.617200000000167, 2.477600000000118]

```
success
Training error : 0.029468289557975657
Weight vector : [29.617200000000167, 2.477600000000118]
```
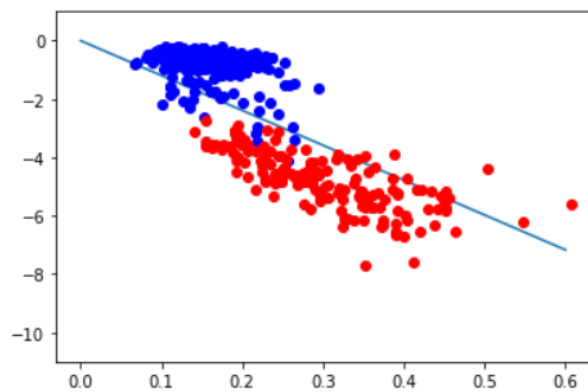

[Training result with optimal weight]

(b), (c) Testing Result

```
error rate : 0.06367924528301887
success
```


[Testing result and Test error rate]

4. Pocket PLA with Raw Feature data

(a)

Algorithm : Pocket

Dataset : X (257*1561), Y(1*1561)

Iteration : Until find optimal weight value which classifing data well

Training error rate : 0.028

Testing error rate : 0.377

This dataset is too large so I iterate 5 times for this whole data. The Training error is kind great, but the test error is bad. I'm looking forward to making better result with bigger iteration.

(b)

```
success                                           error rate :  0.37735849056603776
Training error : 0.028187059577194105             success
```

[Training Result]   [Testing error rate result]

d) **Discussions and Conclusions. Any design decisions you had to make and your experimental observations. What do you observe about the behavior of your program when you run it? Does it seem to work the way you think it should? Play around a little with different setting to see what happens. Note, your open-ended exploration is highly valued.**

For the first dataset, since it is linear separable data, I can easily find the best weight value that classify labels well. After that, for the next dataset, since it is not separable linearly, I should decide which weight value in which iteration is optimal for this model. But the way to choose best iteration is a little confused. I tried to find optimal iteration number through draw a graph with error rate for each iteration in the second non linear dataset, but it seems that error rates are similar each other and I found some seasonality through that graph, and it is not converging.

Also for the last dataset, the test score is quite bad than the training score, which I think it is not that good model, I think, to make this model better, I should scale this data more or do more iteration for this model since I couldn't do much iteration in my computer.

I'm looking forward to making models with Neural Network, or set some batch size and learn the model again, or any other strategies that I want to try. Also, since most of data was not a linear separable data, I want to try other analysis method like making cluster, or nonlinear analyzing method.