# CS534 MACHINE LEARNING

Da Lin 933334812
lind2@oregonstate.edu

## Introduction

In this assignment, students need to build three model: online perceptron, average perceptron, and kernel perceptron. In detail, we need to plot the accuracy performance, find the best value of P, figure out the e asymptotic runtime, and find the best value and iteration in result. At the end, the experiment was fit my expectation. The following section is my answer for all questions.

## Part1:

a. The result shows below, please check the appendix for the value detail of each iteration.

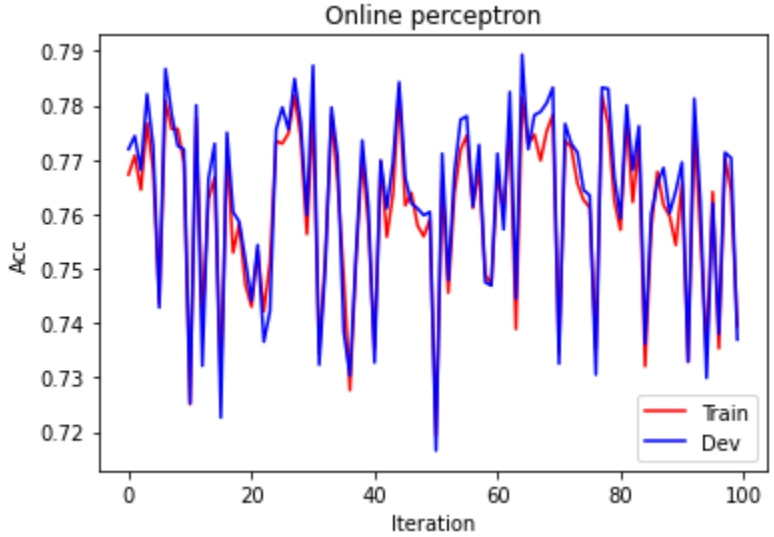


Figure 1.a

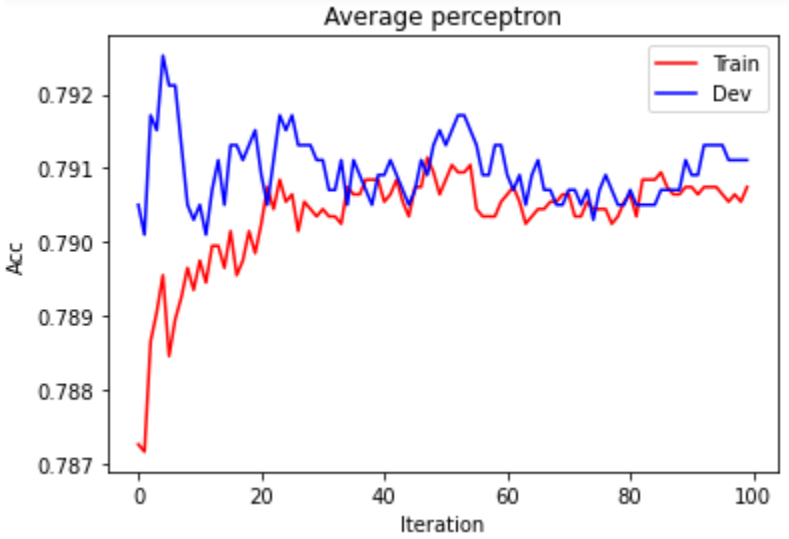

Figure 1.b

b. The performance of average perceptron loos better than the online perceptron, and it did smoother than the online perceptron. I think the reason is the average perceptron takes the average weight of all weights. However, the online perceptron counts all errors in each iteration for updates.

c. The best value is: 0.7925252525252525 on iteration 4

## Part2a:

a & b. The result shows below:

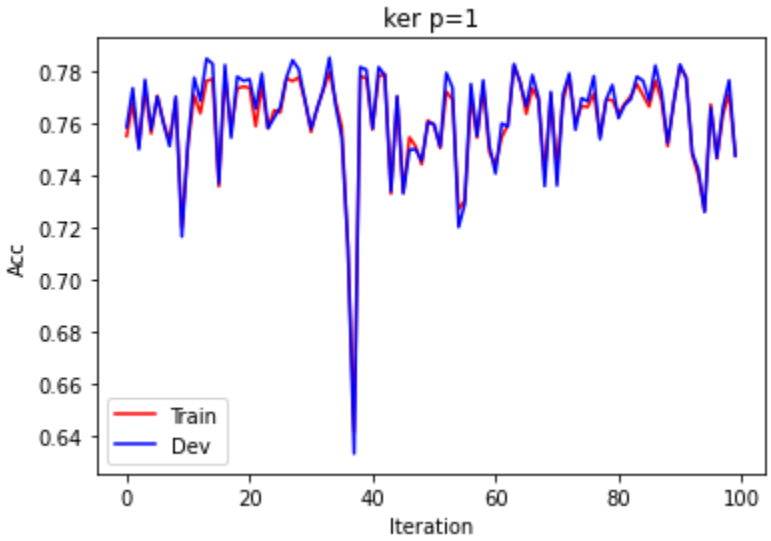


Figure 2a.a

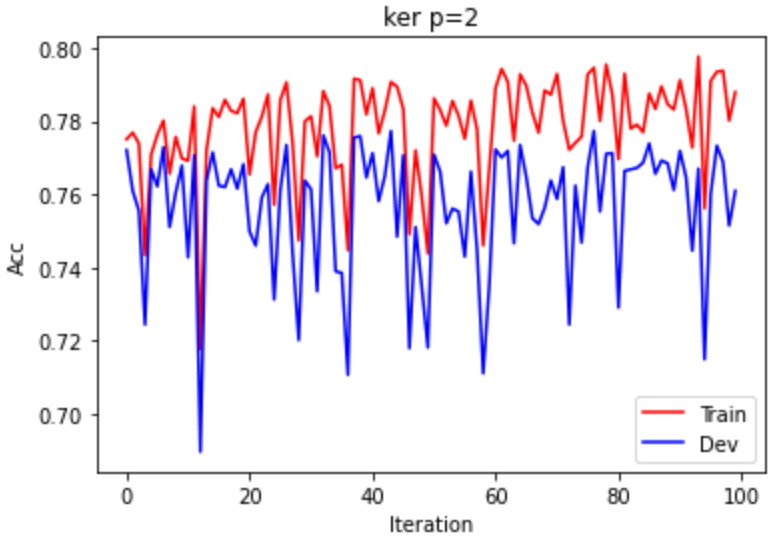


Figure 2a.b

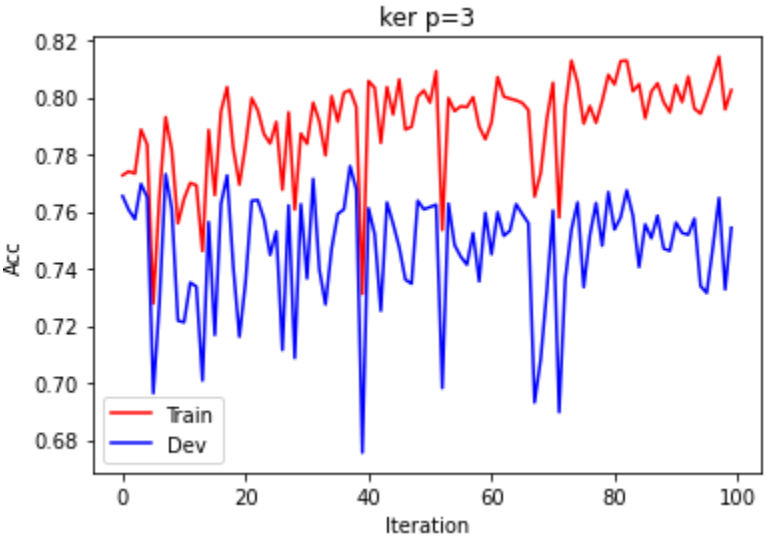


Figure 2a.c

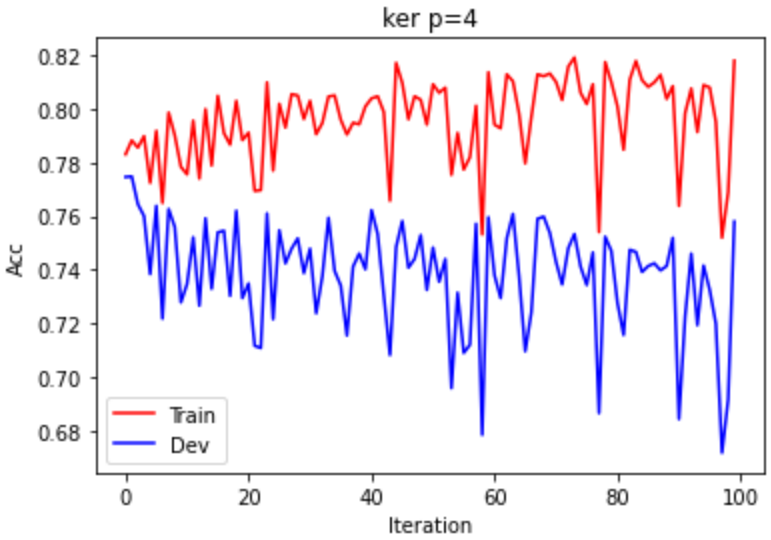


Figure 2a.d

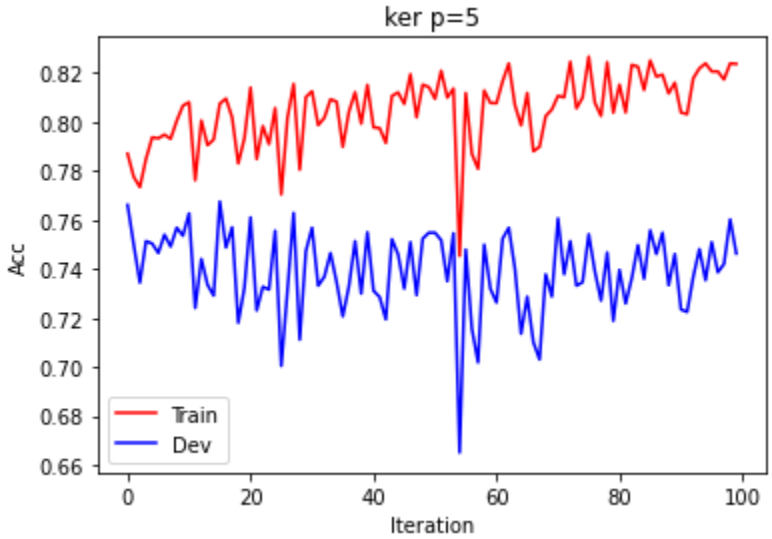


Figure 2a.e

c.

P=1:

The best value is: 0.7852525252525253 on iteration 33.

P=2:

The best value is: 0.7773737373737374 on iteration 43.

P=3:

The best value is: 0.775959595959596 on iteration 37

P=4

The best value is: 0.7747474747474747 on iteration 1.

P=5
The best value is: 0.7674747474747474 on iteration 15.

When P increase, the best value goes down. I think the higher value of P will decrease the accuracy and cause overfitting.

d. The asymptotic runtime of my algorithm is $O(n^2)$. All running time with p value show below:
P=1, The running time: 48.40580463409424
P=2, The running time: 52.41595005989075
P=3, The running time: 52.09683346748352
P=4, The running time: 52.48975372314453
P=5, The running time: 52.41694307327270

## Part2b:

a.

The best value of P is 1. When I change the learning rate, the accuracy performance and learning trajectory did not change. The reason is the data set is not good enough, in this case, when we change the learning rate, the direction of coordinate did not change, it only affects the moving step.

b.

The result shows below:

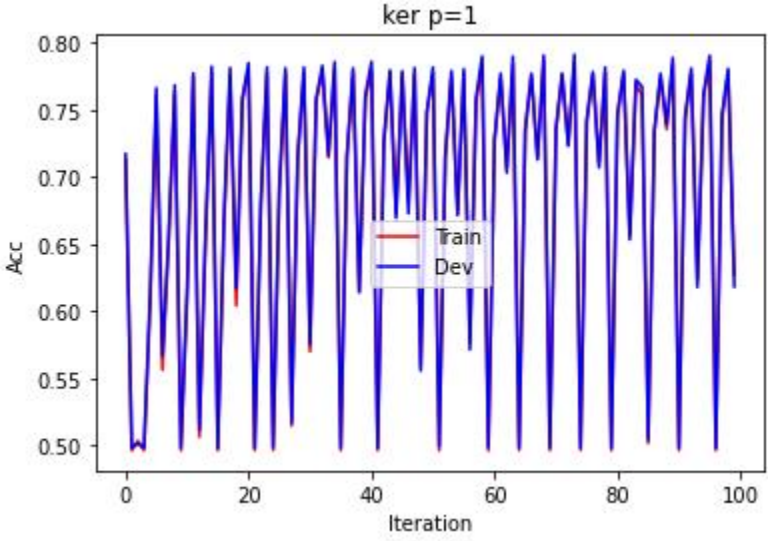


Figure 2b.a

When I add batch, the accuracy performance oscillates a lot than the same value of P in part2a. I think the reason is the data is not linearly separable, so the curve cannot convergence.

c.

The asymptotic runtime of my algorithm is $O(n^2)$.
The running time: 12.911795139312744

The running time decrease a lot because **u** only compute once in each iteration.