

Reinforcement Learning HW-4

Name: Li, Jieda

NetID: jlg7773

Problem 1: Joke recommender using Jester data

Code:

https://github.com/jiedali/reinforcement_learning_jieda_li/blob/master/hw4/hw4_ucb_jester_recommender_final.ipynb

(1) Tuning of hyper-parameter alpha

for each of following alpha value, I further divide the training set into training 80% and validation 20% set, first I fit linear regression on the training set, then for each alpha, compute the average regret on the validation set.

List of alpha:

```
[0.2, 0.5, 0.8, 1, 2, 5, 10, 100, 200, 500, 1000]
```

List of average regret:

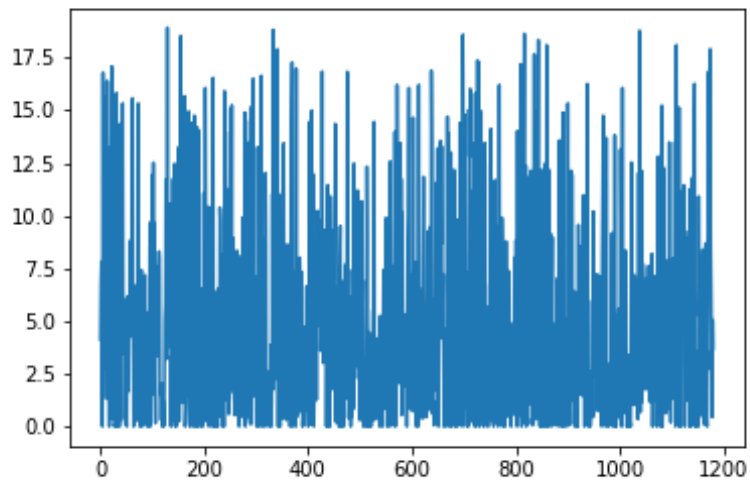
```
[5.027699360933593,  
4.921683801055849,  
4.97693248124479,  
4.936021116976938,  
4.913539872186719,  
5.010266740761322,  
5.005993331480967,  
4.873175882189496,  
4.800752986940817,  
4.945884968046679,  
4.8175715476521255]
```

Alpha=200 gives the minimum regret, we will set alpha = 200

(2) With alpha=200, plot the regret for the remaining 1181 persons

Now we set alpha=200, fit the linear model on the entire training set 18000 rows and compute the regret for the remaining 1181 persons

Following plot shows the regret for each person



Average regret is: 4.881574936494496

Maximum regret: 18.93

Minimum regret: 0.0