COMS 6998-2 Advanced Machine Learning with Personalization (Spring 2018)
Assignment #2

Shreyas Mundhra - ssm2211@columbia.edu

March 25, 2018

1 Introduction

The aim of this assignment is to implement and evaluate the linUCB algorithm to recommend news articles in an online manner. For the purpose of this assignment, the news articles (arms) are assumed to be constant for all the trials and there are a total of 10 news articles. The first dataset (dataset.txt) was used for the purpose of this assignment. This dataset consists of 10000 trials. In each trial, information about the recommended news article for that trial, the corresponding reward for that article and the context for that trial consisting of 100 features is provided. The reward is 1 if a news article is clicked, and 0 otherwise.

2 Implementation

The (disjoint) linUCB algorithm was implemented for this assignment. However, we only know the actual reward for the arm provided for that trial in the dataset and the arm that has the highest estimated payoff according to the algorithm in that trial might not be the same as the arm that is actually chosen. Hence, we only update the parameters (A_{a_t} and b_{a_t} for some a_t) for the arm chosen in that trial in the dataset and not the arm having the highest estimated payoff according to the algorithm for that trial.

At the same time, we keep evaluating the performance of the algorithm by computing the cumulative take-rate (CTR) in an online manner. We do this by keeping track of the trials that predict the same arm that is provided in the dataset for that trial. Since we do this in an online manner, the predicted arm for time t is the arm having the highest estimated payoff using the algorithm trained up to time $t-1$. We then evaluate the CTR over these trials using the following formula:

$$C(T) = \frac{\Sigma_{t=1}^{T} y_t * \mathbb{1}[\Pi_{t-1}(x_t) = a_t]}{\Sigma_{t=1}^{T} \mathbb{1}[\Pi_{t-1}(x_t) = a_t]} \tag{1}$$

Here, Π_{t-1} is the algorithm trained on data up to time $t-1$. $\Pi_{t-1}(x_t)$ is the action that algorithm Π_{t-1} chooses for the context x_t which it observes at time t. a_t is the real action that was taken in the data-set at time t. y_t is the real reward that was obtained at time t.

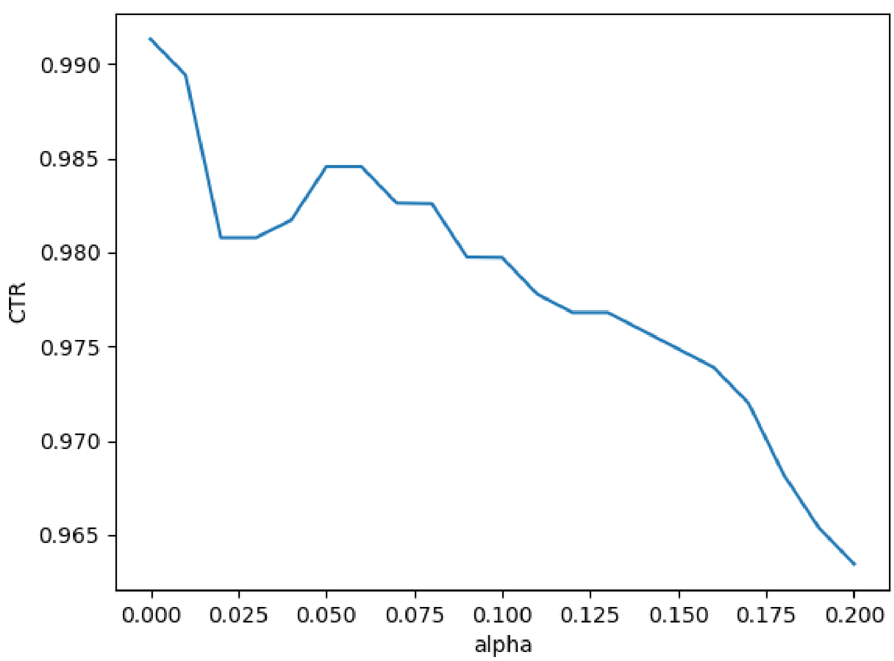
3 α tuning

Three different strategies were tried for setting α as explained below.

3.1 Strategy 1

In this strategy, I fixed α to a constant value throughout all the trials.

CTR was calculated for different values of α from 0 to 0.2 (both inclusive) with increments of 0.01. The plot of CTR versus α for different values of α is as shown below:



The value of CTR for each setting of α is given in the file "./data/ctrVsAlpha/Strategy1.txt".

We can see that a maximum CTR of 0.9913294797687862 was obtained for $\alpha = 0$.

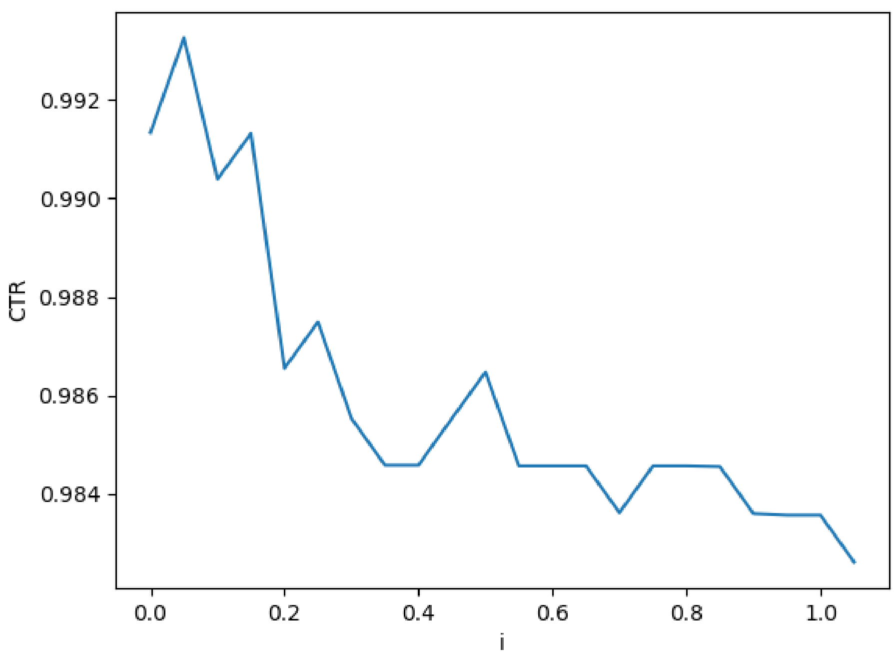
3.2 Strategy 2

In this strategy, I varied α based on the trial number t as follows:

$$\alpha = \frac{i}{\sqrt{t}} \tag{2}$$

for some i.

CTR was calculated for different values of i from 0 to 1 (both inclusive) with increments of 0.05. The plot of CTR versus i for different values of i (and consequently α for a given trial t) is as shown below:



The value of CTR for each setting of α is given in the file "./data/ctrVsAlpha/Strategy2.txt".

We can see that a maximum CTR of 0.9932562620423893 was obtained for $i = 0.05$.

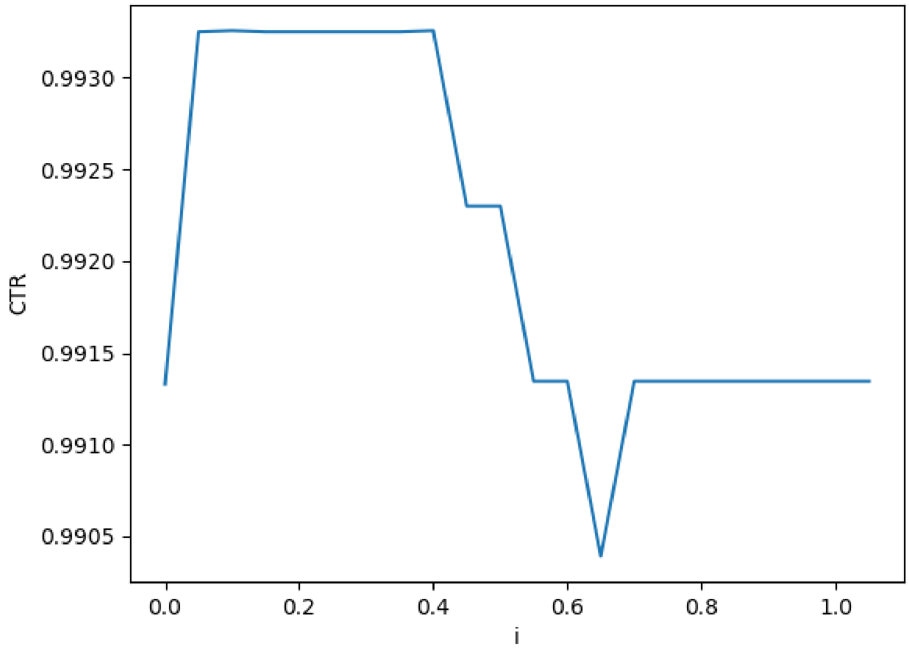
3.3 Strategy 3

In this strategy, I varied α based on the trial number t as follows:

$$\alpha = \frac{i}{t} \tag{3}$$

for some i.

CTR was calculated for different values of i from 0 to 1 (both inclusive) with increments of 0.05. The plot of CTR versus i for different values of i (and consequently α for a given trial t) is as shown below:



The value of CTR for each setting of α is given in the file "./data/ctrVsAlpha/Strategy3.txt".

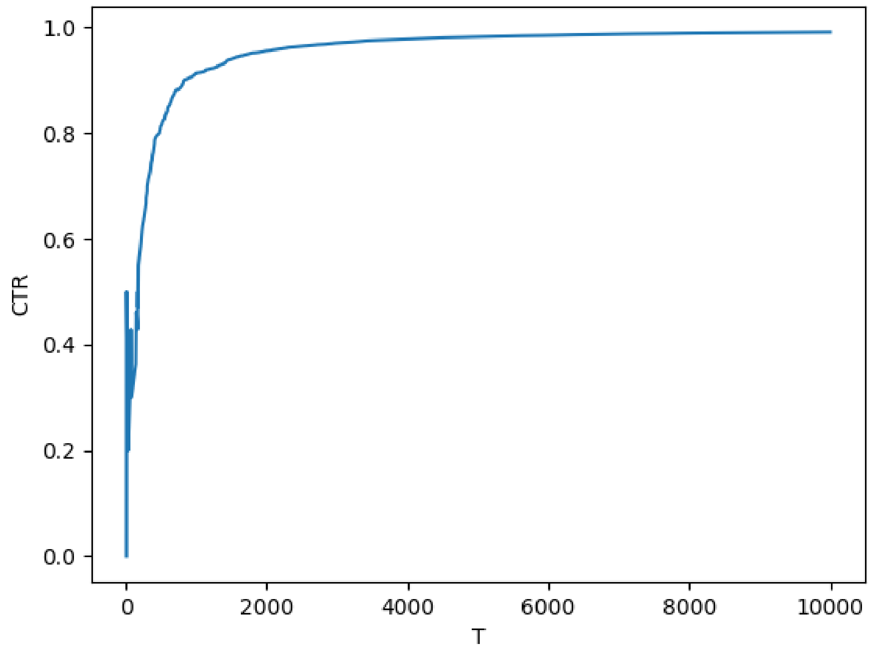
Although from the plot, it seems as if the maximum CTR was obtained for all values of i between 0,05 and 0.4 (both inclusive), the CTR was slightly higher for 0.1 and 0.4, which was 0.9932562620423893, compared to the CTR of 0.9932497589199615 for the other values of i between 0.05 and 0.4.

4 Results

The plots for the CTR versus trial T were plotted for the best setting of α found for each setting.

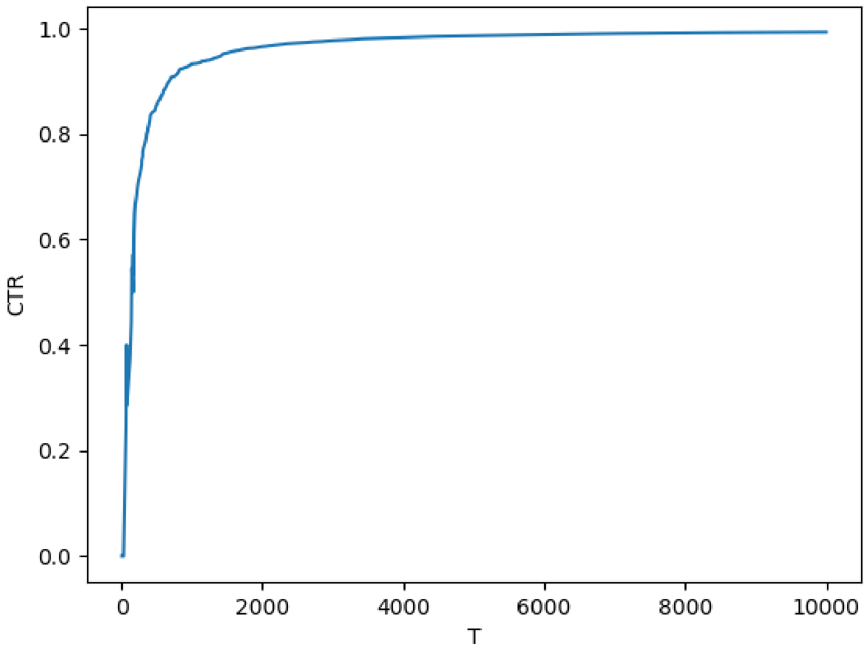
4.1 Strategy 1

As discussed in the previous section, the best setting for this strategy was $\alpha = 0$ for which a CTR of 0.9913294797687862 was obtained. The plot is as shown below.



4.2 Strategy 2

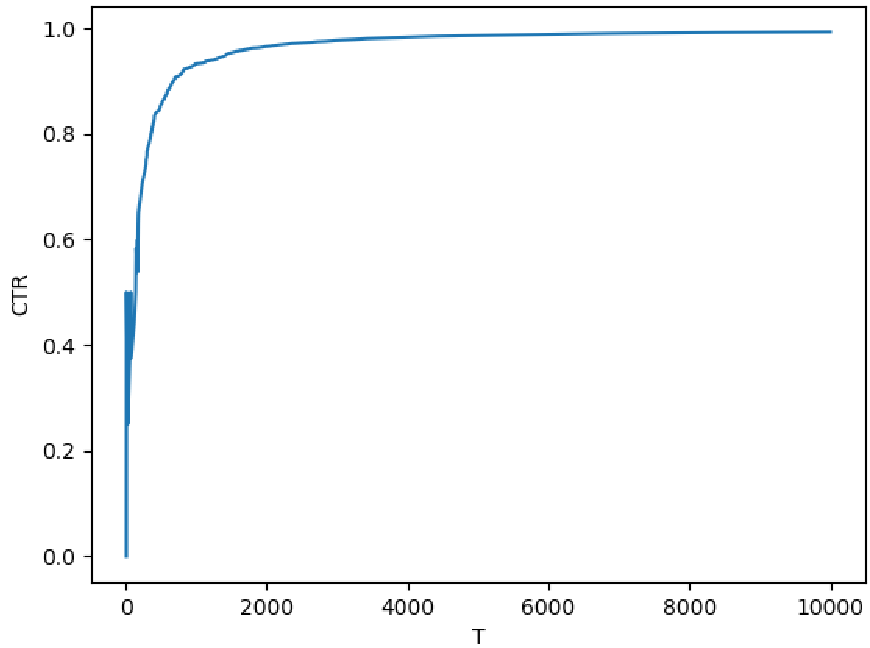
As discussed in the previous section, the best setting for this strategy was $\alpha = \frac{0.05}{\sqrt{t}}$ (for $i = 0.05$) for which a CTR of 0.9932562620423893 was obtained. The plot is as shown below.



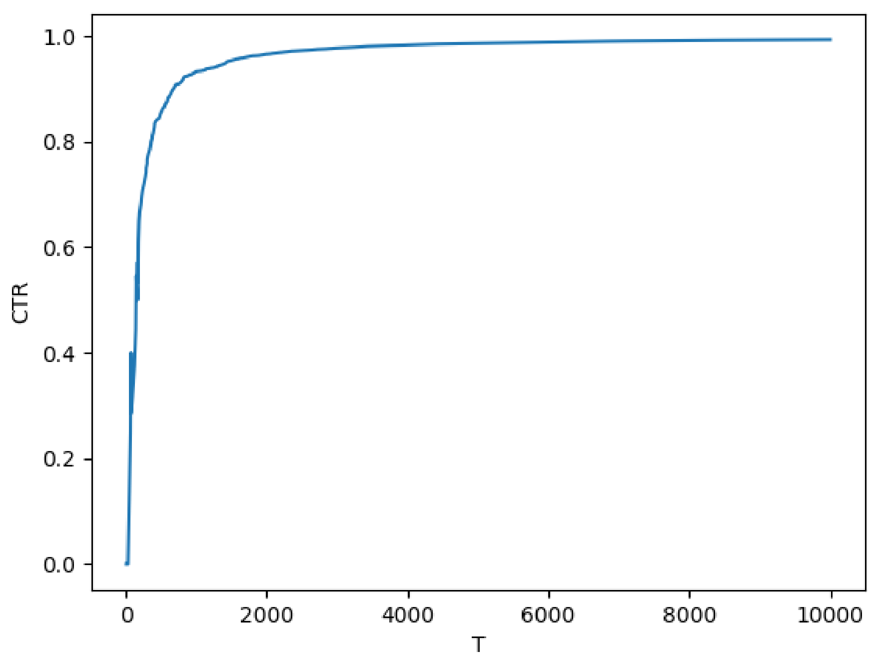
4.3 Strategy 3

As discussed in the previous section, the best settings for this strategy were $\alpha = \frac{0.1}{t}$ (for $i = 0.1$) and $\alpha = \frac{0.4}{t}$ (for $i = 0.4$) for which a CTR of 0.993256262042389 was obtained.

The plot for $\alpha = \frac{0.1}{t}$ is as shown below.



The plot for $\alpha = \frac{0.4}{t}$ is as shown below.



4.4 Best α Settings

From all the above plots, we can conclude that the best performance can be obtained by setting $\alpha = \frac{0.05}{\sqrt{t}}$, $\alpha = \frac{0.1}{t}$ or $\alpha = \frac{0.4}{t}$, all of which give a CTR of 0.993256262042389.