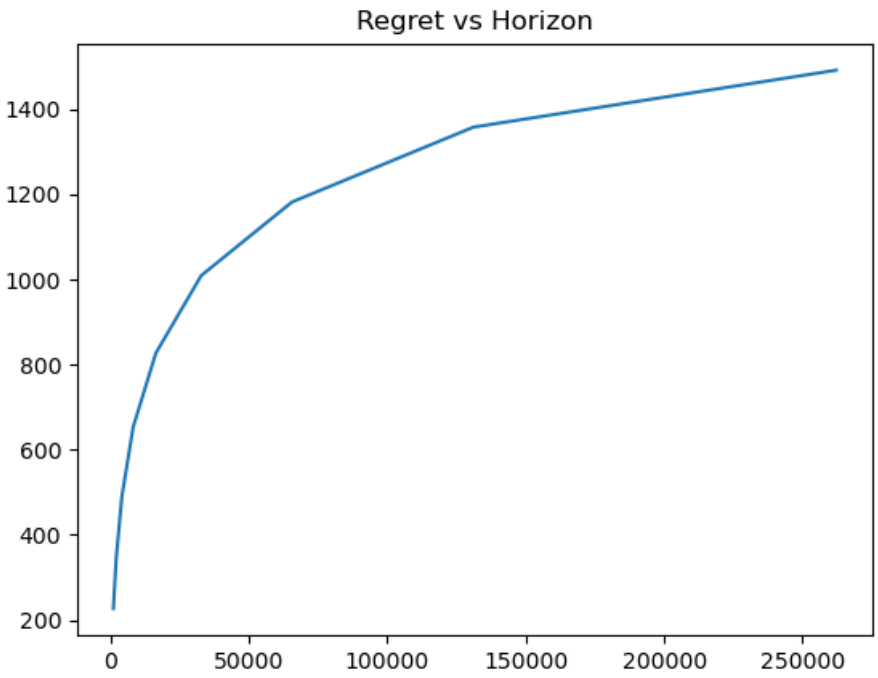
Satwik Murarka
190020101

# Programming Assignment 1: CS747

## Task 1

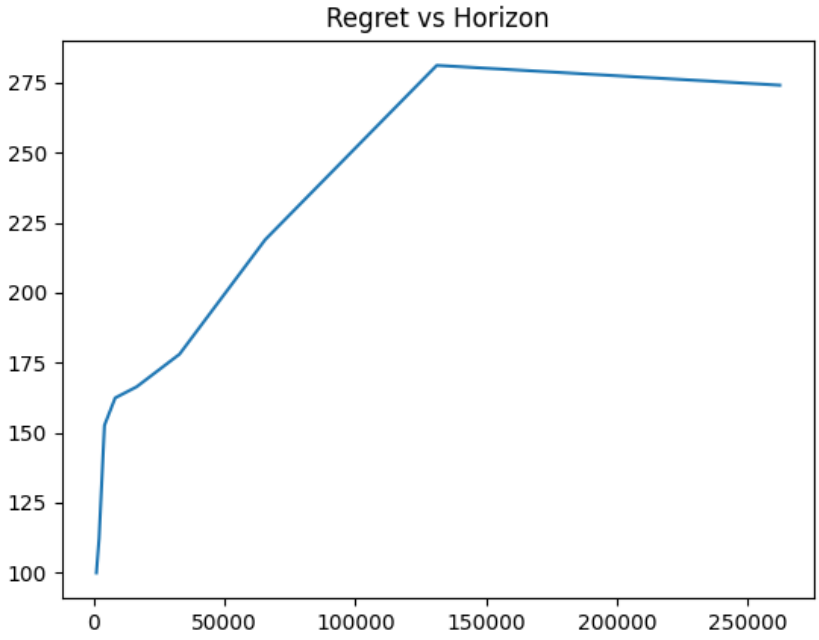
**UCB Algorithm:**




**Plot 1: Regret vs Horizon for UCB Algorithm**

The following is the plot for regret vs horizon for the UCB algorithm. For the implementation, I sampled each arm once so that the UCB is defined. Post that I took the arm with the maximum UCB. At each time step, I updated the UCB for each arm and number of pulls, and the empirical mean for the pulled arm.
The plot seems to be below the linear (y=x) line for large horizons which is expected because of the sub-linear (logarithmic) behavior of the algorithm.

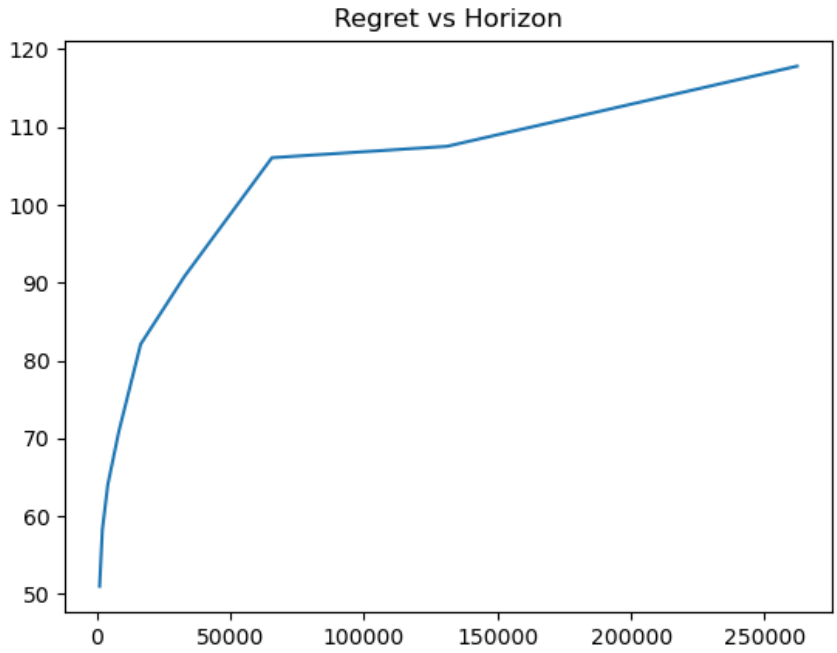
**KL-UCB Algorithm:**




**Plot 2: Regret vs Horizon for KL-UCB Algorithm**

The following is the plot for regret vs horizon for the KL-UCB algorithm. For the implementation, I sampled each arm once so that the KL-UCB is defined. Post that I used a binary search in the interval of empirical mean and 0.99 (with an interval of 0.01) to find the maximum ucb-kl for each arm. Following this, I the arm with the maximum kl-ucb was pulled. At each time step, I updated the kl-ucb for each arm and number of pulls, and the counts for the pulled arm.
The plot seems to be below the linear (y=x) line for large horizons which is expected because of the sub-linear(logarithmic) behavior of the algorithm.

**Thompson Sampling Algorithm:**




**Plot 3: Regret vs Horizon for Thompson Sampling Algorithm**

The following is the plot for regret vs horizon for the Thompson Sampling algorithm. For the implementation, I noted the success and failure of each arm and used them for the parameters of the beta distribution for each arm. From the beta distribution of each arm, a sample is sampled and this is compared across the multiple arms with the arm with the highest sample being selected. At each time step, I updated the success and failure for each arm.
The plot seems to be below the linear (y=x) line for large horizons which is expected because of the sub-linear(logarithmic) behavior of the algorithm.

# Task 2

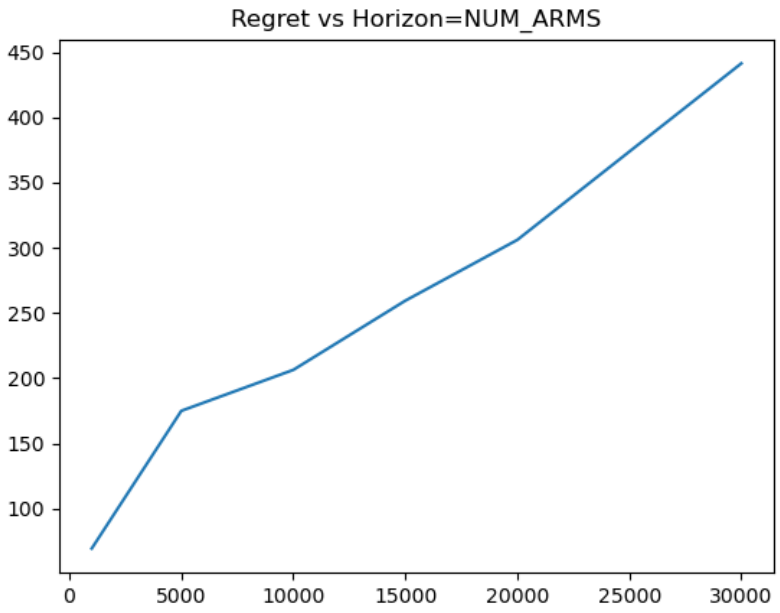

**Plot 4: Regret vs Horizon for Batched Sampling**

The following is the plot for regret vs horizon for the case we sample the arms in batches . For the implementation, I used the GLIE version of the epsilon first algorithm.

This is done by exploring by sampling uniformly till batches which are less than $\sqrt{Horizon}$ and then choosing the arm with the highest empirical mean for all the samples of the batch.

The plot seems to be growing with the linear (y=x) line which implies that the algorithm achieves linear regret.

# Task 3




**Plot 4: Regret vs Horizon for Horizon = Number of Arms**

The following is the plot for regret vs horizon for the case when the Horizon is equal to the number of arms. For the implementation, I used the GLIE version of the epsilon first algorithm. This is done by exploring by sampling uniformly till $\sqrt{Horizon}$ and then choosing the arm with the highest empirical mean with updating it in each time step.
The plot seems to be below the linear (y=x) line which implies that the algorithm achieves sub linear regret.