CS 747: Programming Assignment 1

Aman Singh : 190020010

Task 1

Assumptions

We resolve ties in emperical means in Epsilon Greedy algorithm, ucb values in UCB algorithm, kl-ucb values in KL-UCB algorithm and sampled values in Thompson Sampling algorithm by taking smallest index arm. Also, we pull all arms in the start in all but the Thompson Algorithm. We use c=3 in KL-UCB.

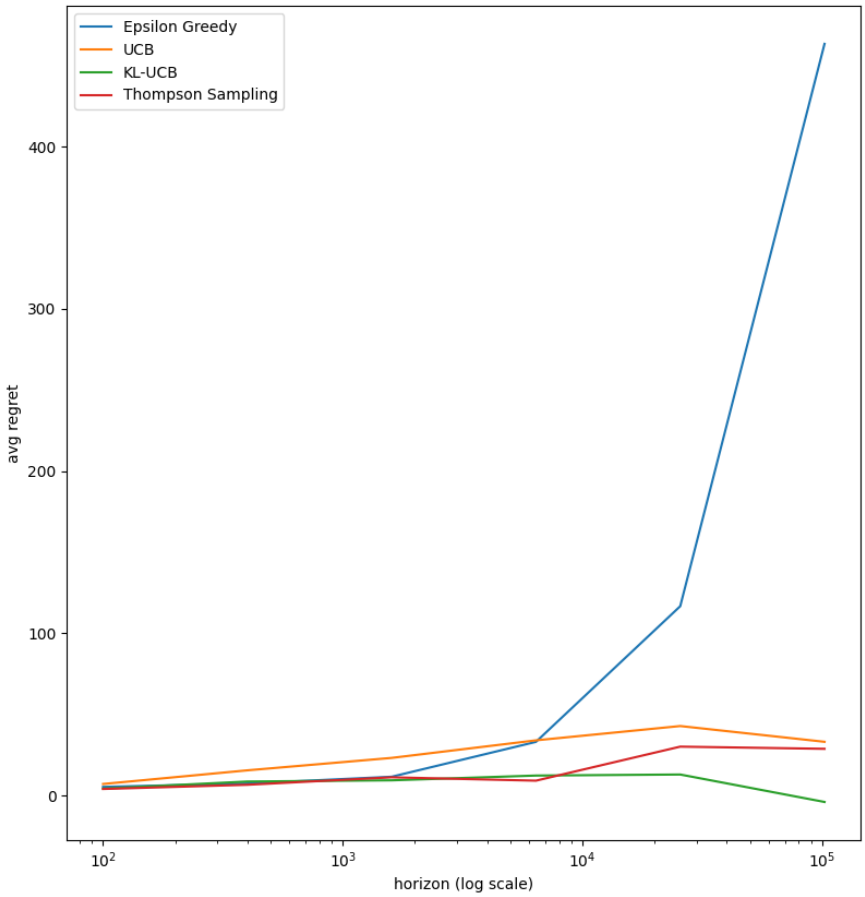


Figure 1: Instance 1

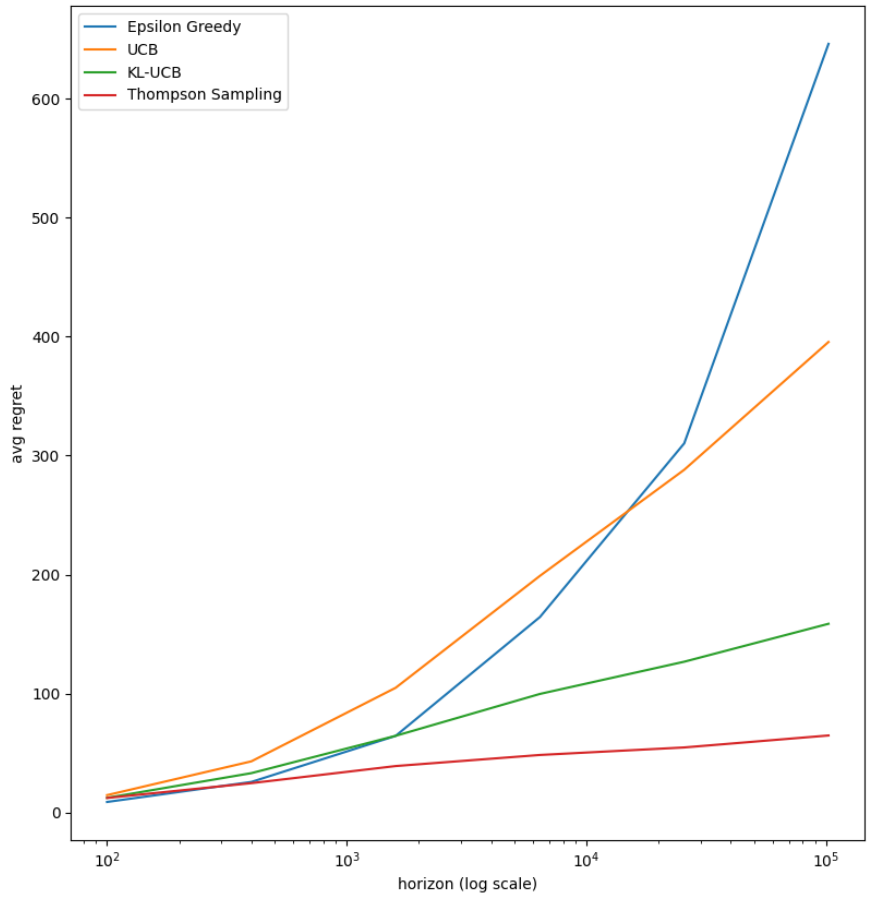


Figure 2: Instance 2

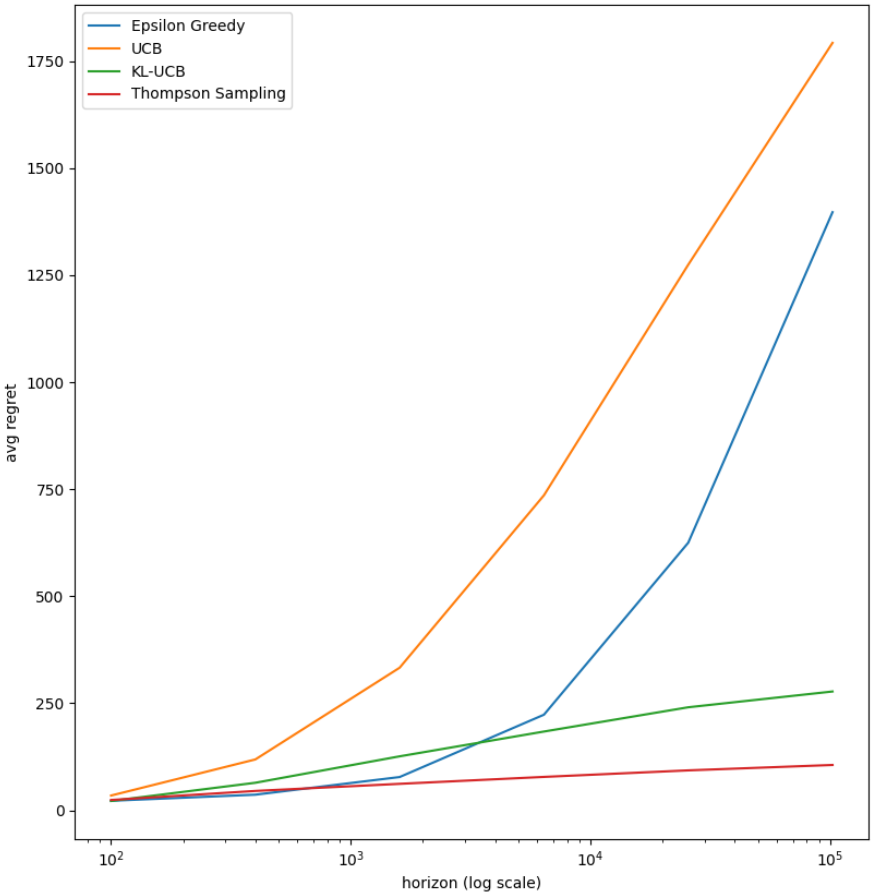
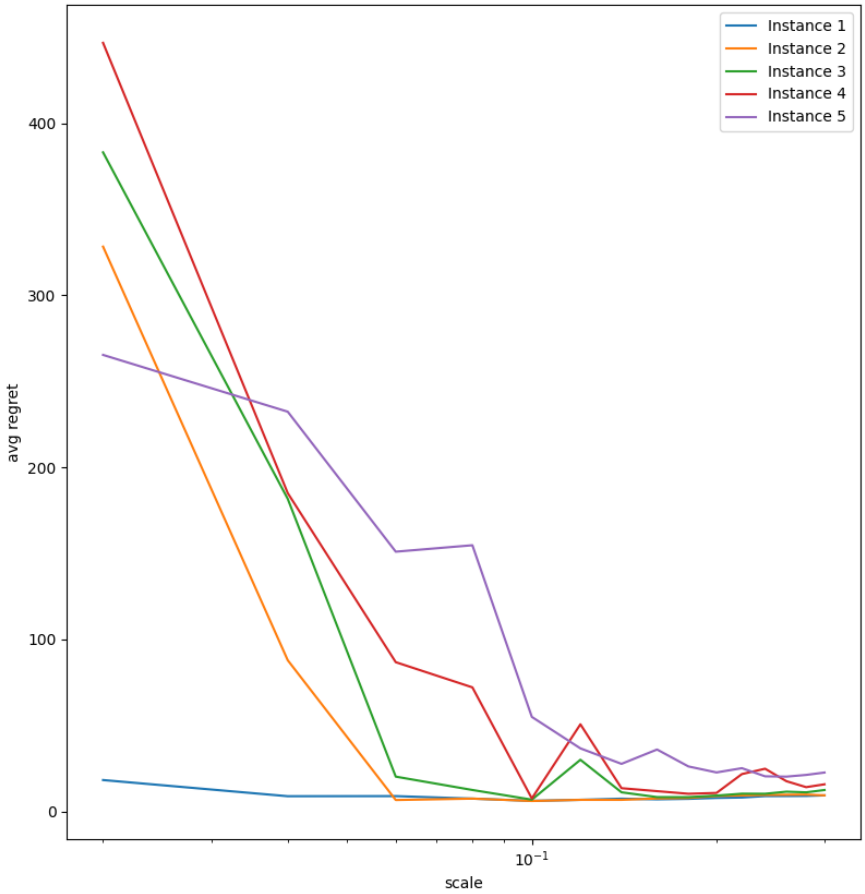


Figure 3: Instance 1

1 Task 2

Following are the values of c for minimum regret :
`Instance 1` : 0.04
`Instance 2` : 0.06
`Instance 3` : 0.10
`Instance 4` : 0.10
`Instance 5` : 0.24
The regret values decrease as we increase scale because scale factor defines how much exploration an algorithm performs and sufficient exploration is important to minimize regret.



2 Task 3

Thompson Sampling algorithm is GLIE and achieves minimum possible regret. Also from Task 1, we have good confidence in it. So we implement the algorithm to solve part 3 such that when we sample an arm and say get 'x' we add (x) to success term for the arm and (1-x) to failure term.

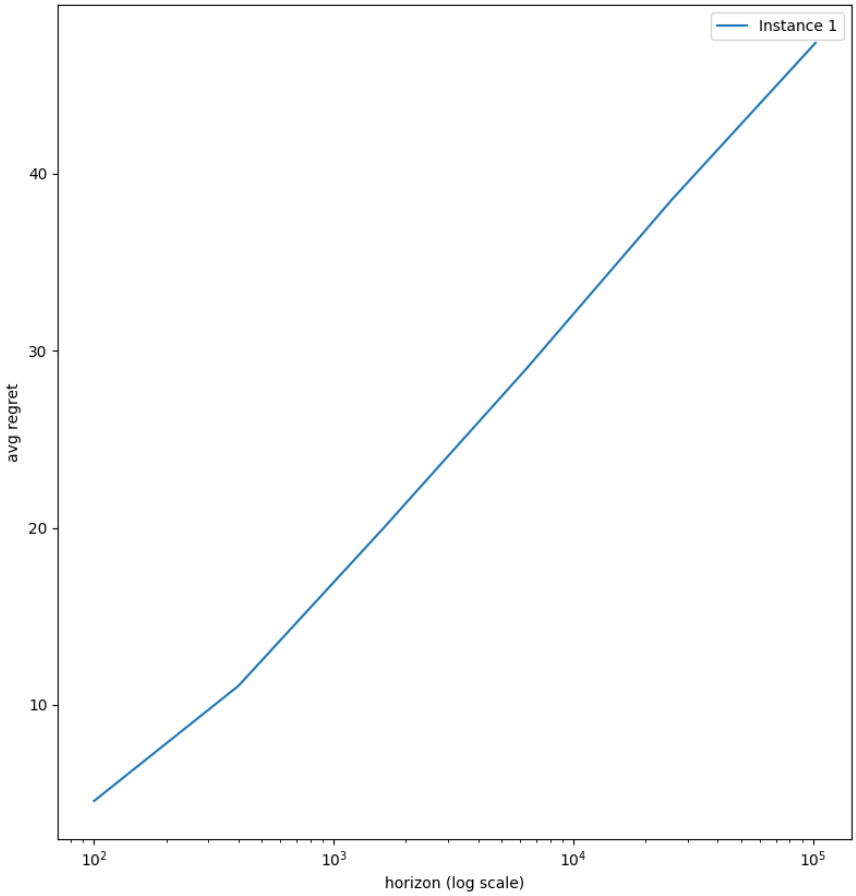


Figure 4: Instance 1

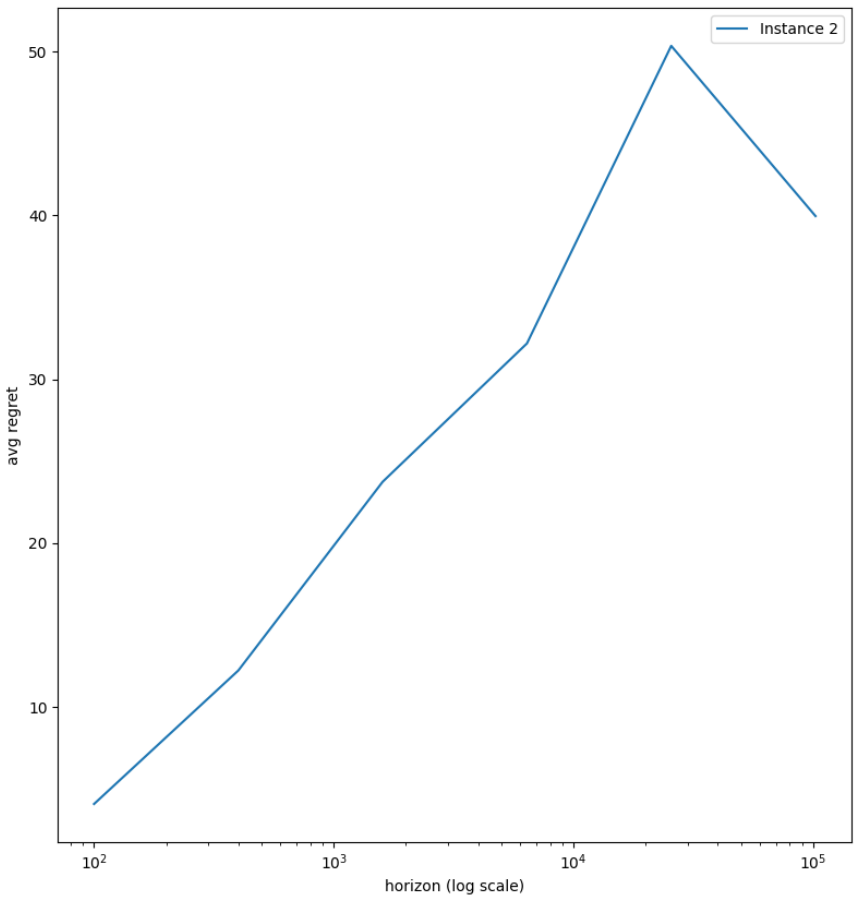


Figure 5: Instance 2

3 Task 4

We claim that Task 4 is exactly the same as Task 1. We need to maximize HIGHS. The probability that you get a high from pulling an arm is (roughly speaking) just $\left(\sum_{x>threshold} \mathbb{P}(x)\right)$ and $\left(1 - \sum_{x>\text{threshold}} \mathbb{P}(x)\right)$. This is just the same as the bernoulli arm (support = {0,1}) in Task 1. So we just use the best algorithm from Task 1 i.e. Thompson Sampling, we add 1 to the number of successes if we get HIGH else to the number of failiures.

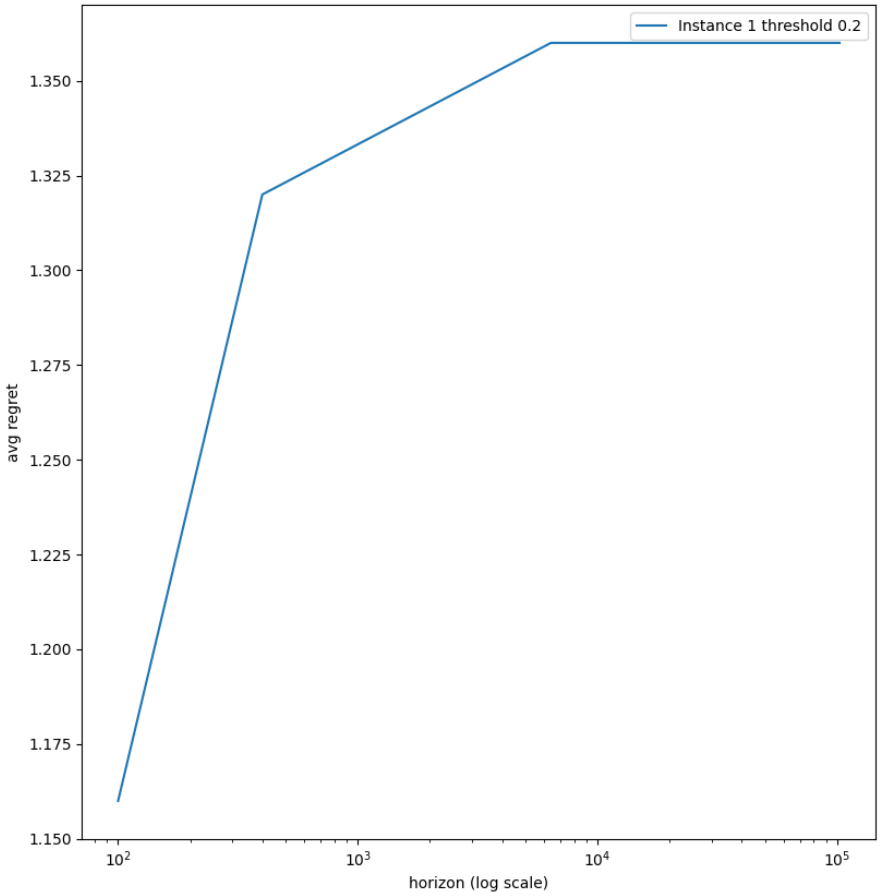


Figure 6: Instance 1 Threshold 0.2

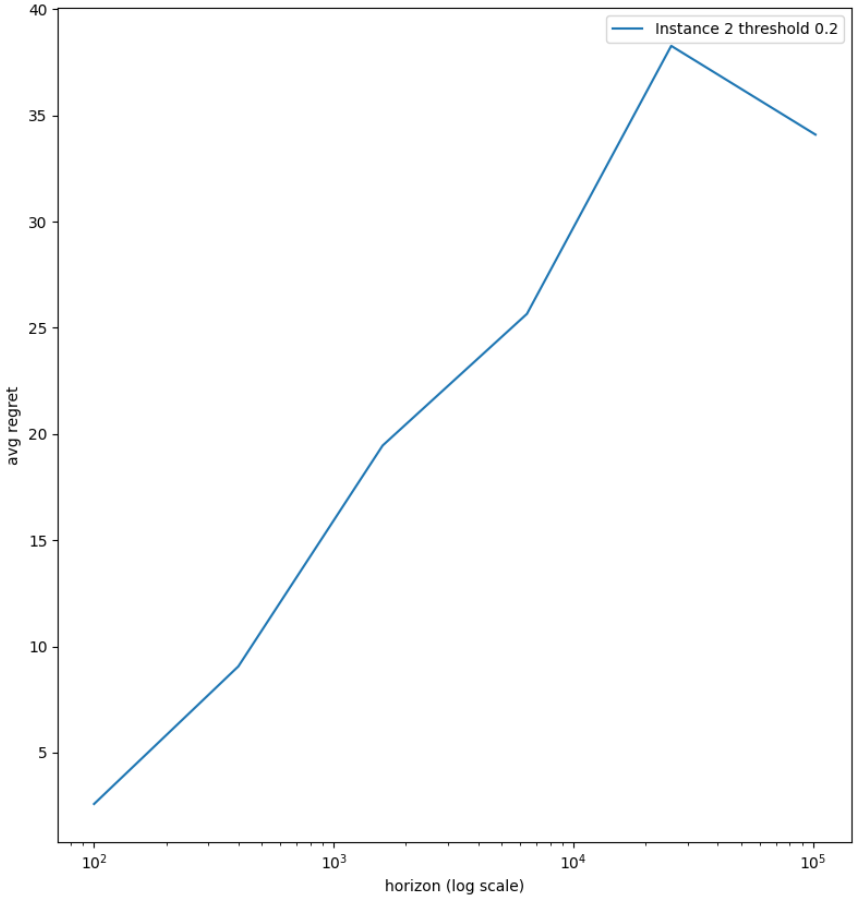


Figure 7: Instance 2 Threshold 0.2

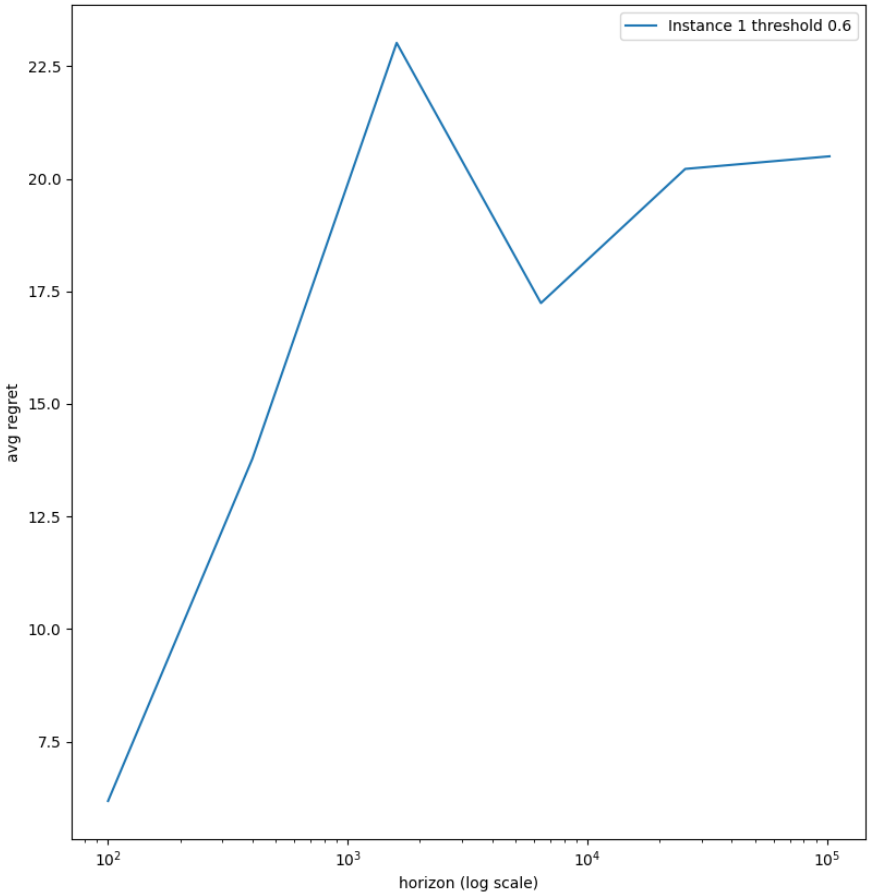


Figure 8: Instance 1 Threshold 0.6

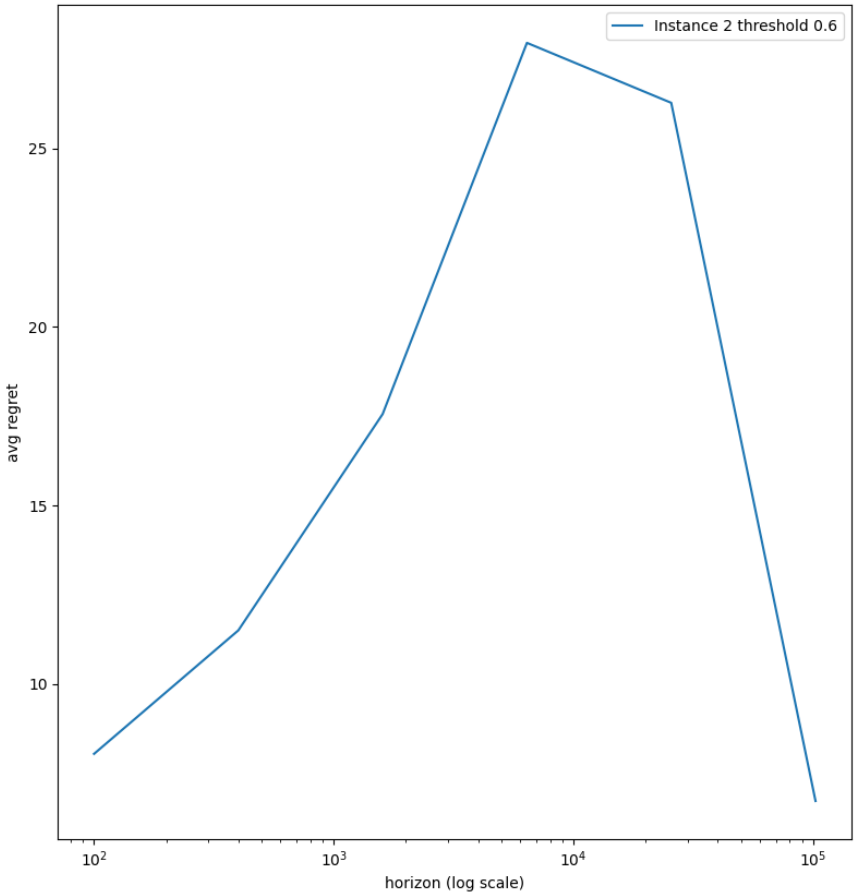


Figure 9: Instance 2 Threshold 0.2