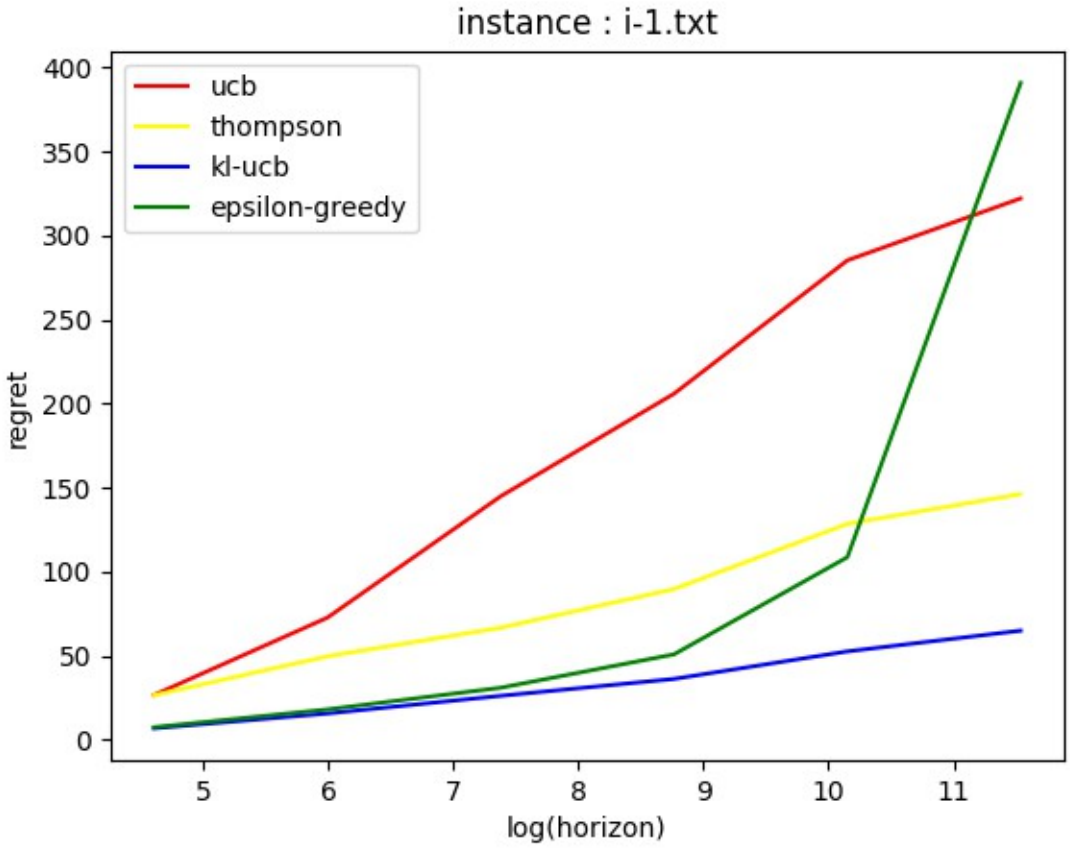
FILA ----PA01 170050053---SUDHIR KUMAR

task1----- graphs

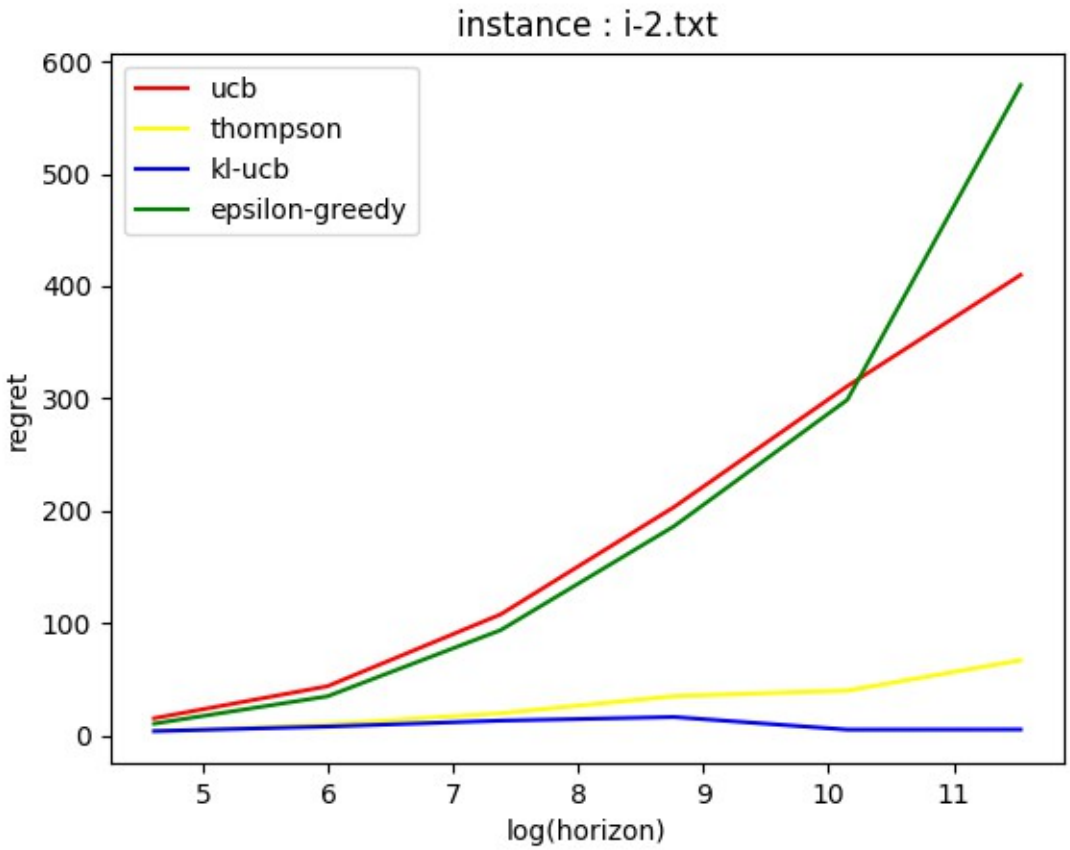


assumption:

1. first of all pull each arm atleast once

observation:

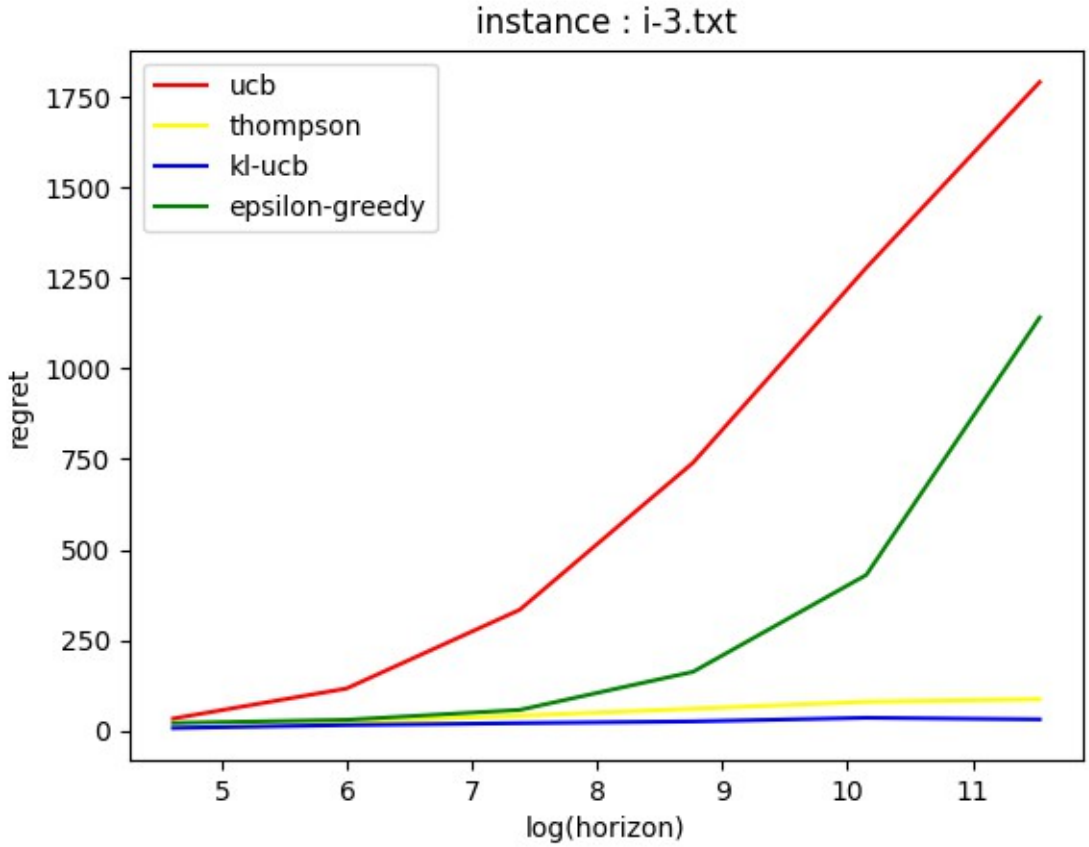
kl-ucb is giving lesser regret as compared to other algorithms
thompson sampling is giving better results for larger horizon
epsilon greedy is not bad choice as compared to ucb for small horizon



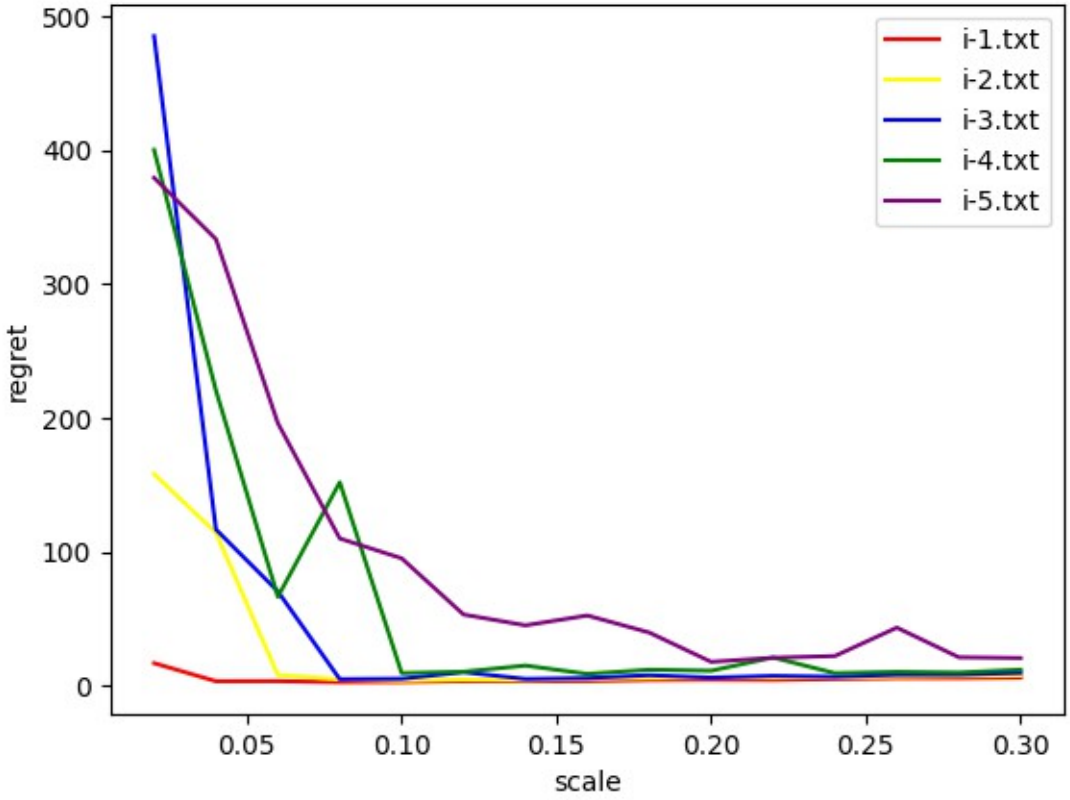
approach:

created algorithms (ucb, kl-ucb and other)in separate file
each algorithm file have class method with required attributes and functions
all these functions are imported in bandit.py
bandit.py takes arguments as per given
created wrapper bash file to generate data
using data plotted graphs

similar for all algorithms (tried to comment in code itself):
when called resume_untill(horizon):
select pull_A : select arm to to pull according to same greedy
ret-reward : returns rewards for particular arm
update empirical mean for time step
update number of pulls
update regret


instance : i-3.txt
ucb
thompson
kl-ucb
epsilon-greedy
regret
log(horizon)
0
250
500
750
1000
1250
1500
1750
5
6
7
8
9
10
11

task2---- graph



approach:

executed wrapper bash script to run bandit.py file

which imports ucb-t2 file contents and

and assoiciated functions

description of functions are commented in files itself

returned vlues are redirected to outputdata.txt by wrapper

observation

larger the value of scale c lower the regret