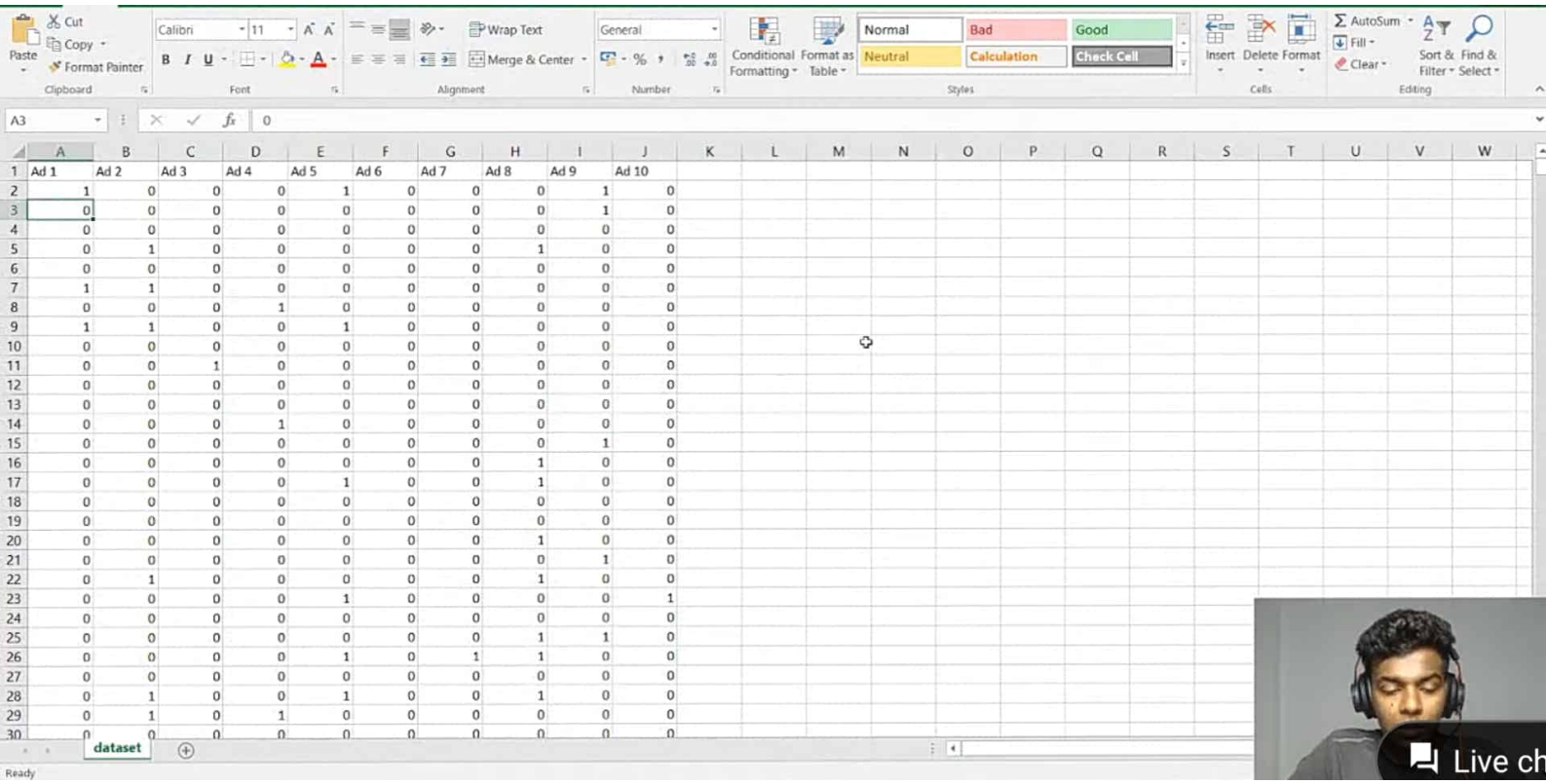
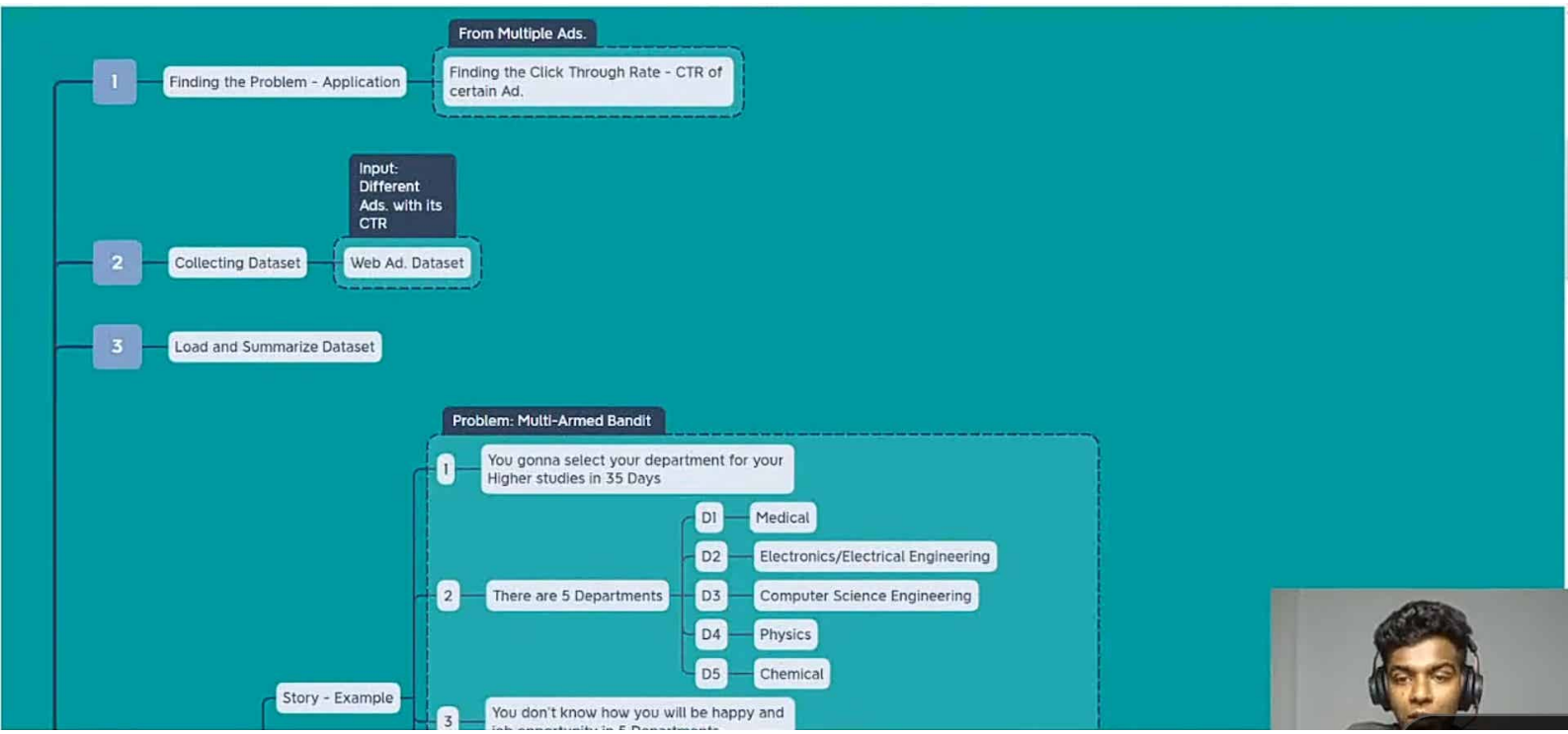
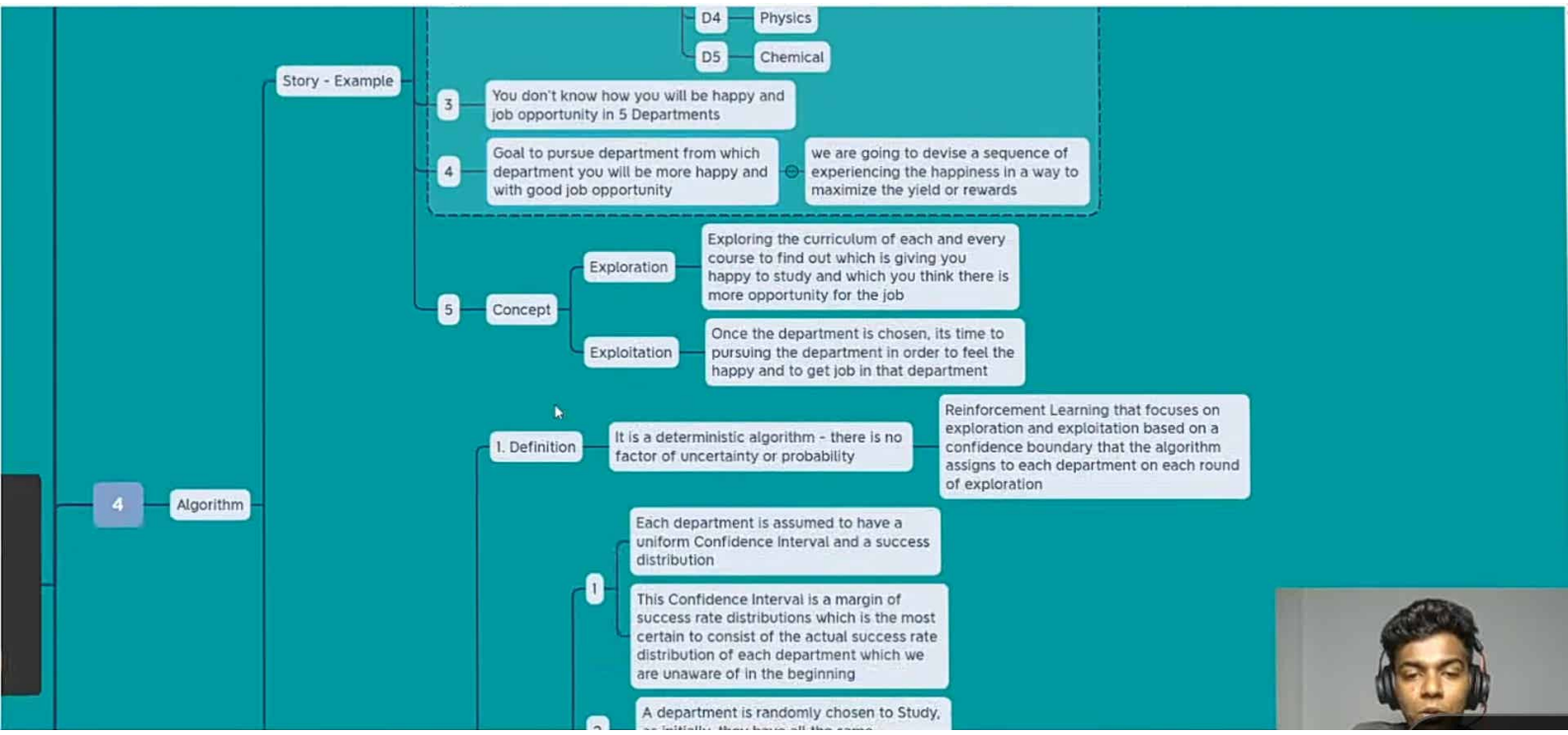
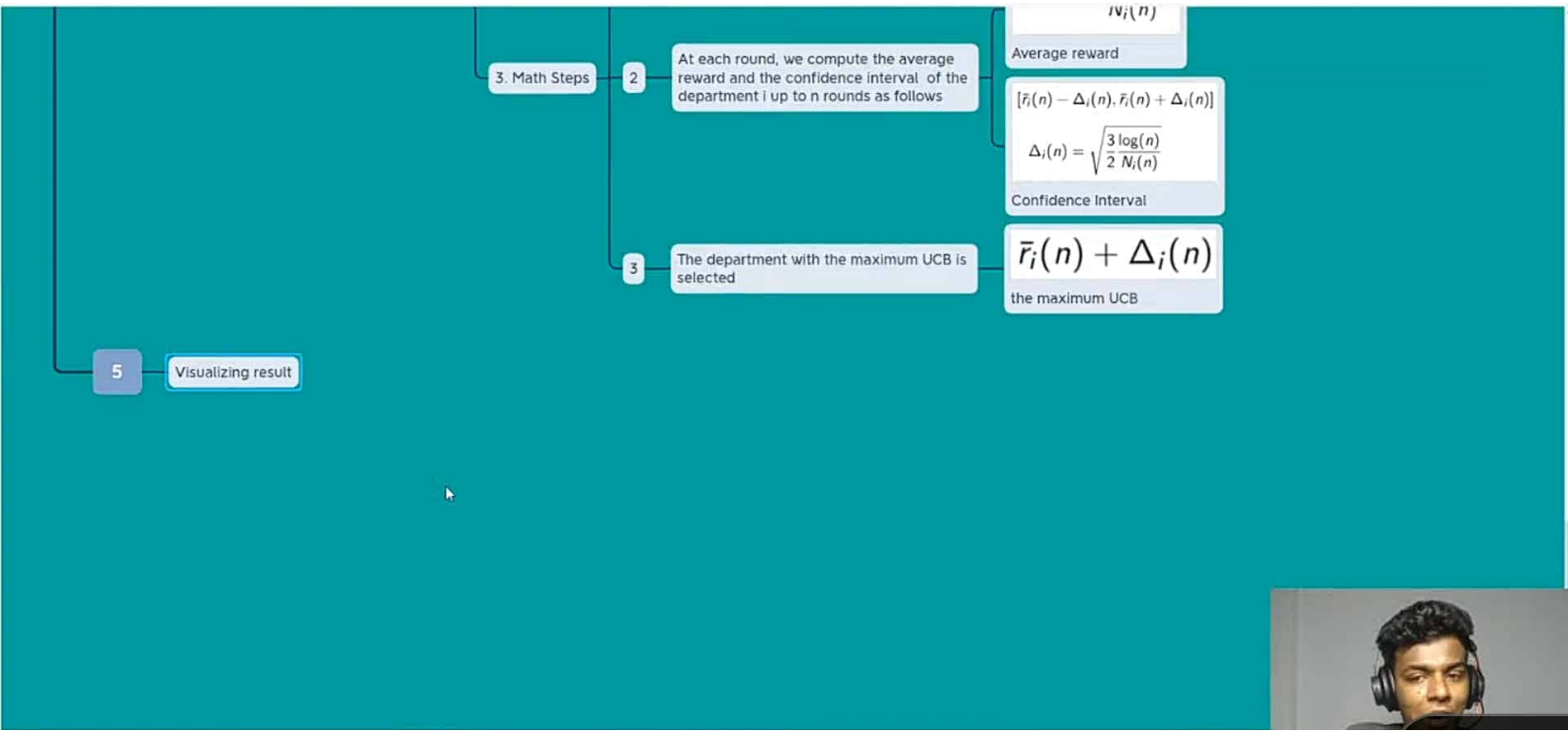


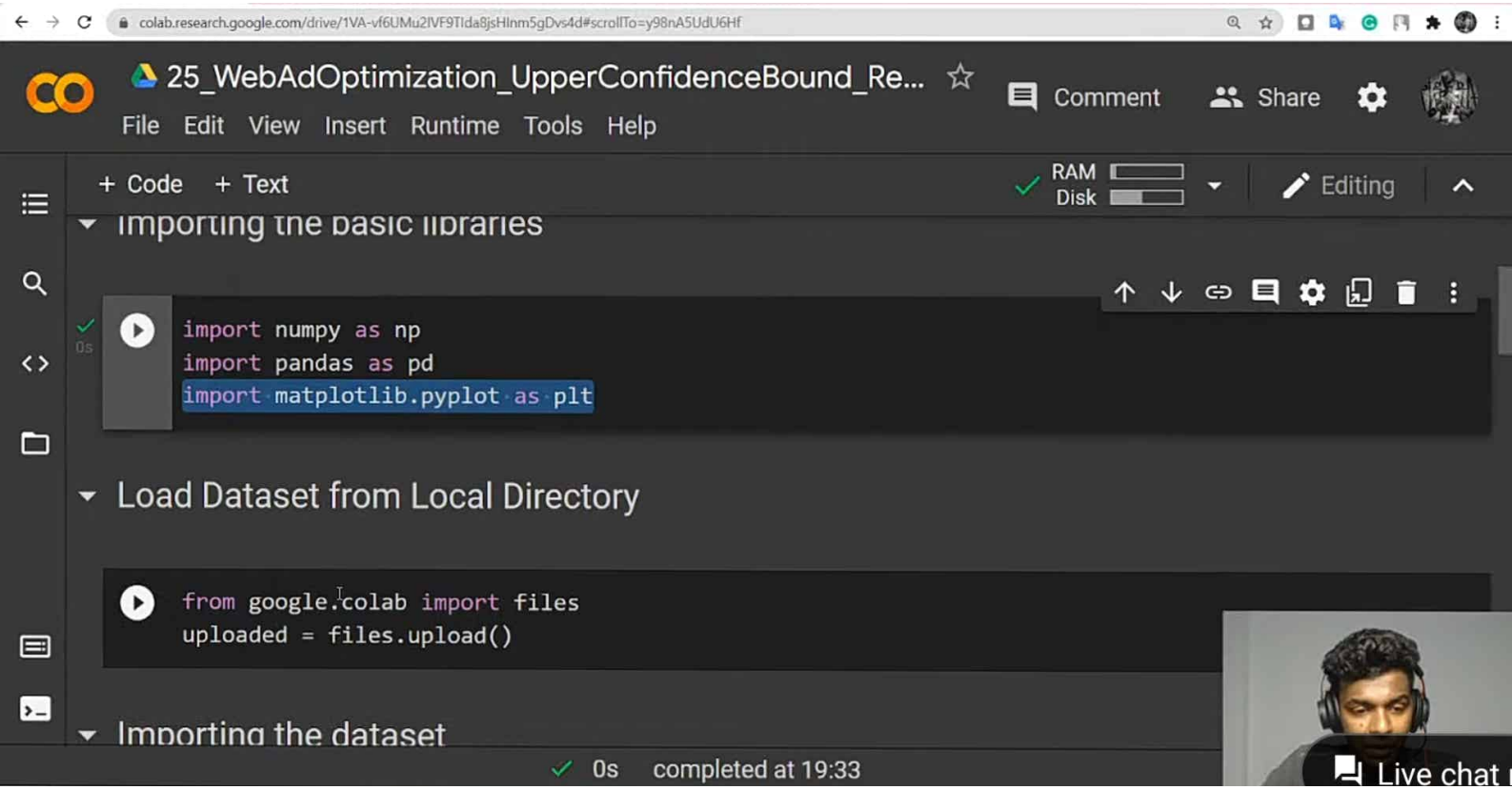
	A	B	C	D	E	F	G	H	I	J
1	Ad 1	Ad 2	Ad 3	Ad 4	Ad 5	Ad 6	Ad 7	Ad 8	Ad 9	Ad 10
2	1	0	0	0	1	0	0	0	1	0
3	0	0	0	0	0	0	0	0	1	0
4	0	0	0	0	0	0	0	0	0	0
5	0	1	0	0	0	0	0	1	0	0
6	0	0	0	0	0	0	0	0	0	0
7	1	1	0	0	0	0	0	0	0	0
8	0	0	0	1	0	0	0	0	0	0
9	1	1	0	0	1	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0
11	0	0	1	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0
14	0	0	0	1	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	1	0
16	0	0	0	0	0	0	0	1	0	0
17	0	0	0	0	1	0	0	1	0	0
18	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	1	0	0
21	0	0	0	0	0	0	0	0	1	0
22	0	1	0	0	0	0	0	1	0	0
23	0	0	0	0	1	0	0	0	0	1
24	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	1	1	0
26	0	0	0	0	1	0	1	1	0	0
27	0	0	0	0	0	0	0	0	0	0
28	0	1	0	0	1	0	0	1	0	0
29	0	1	0	1	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0

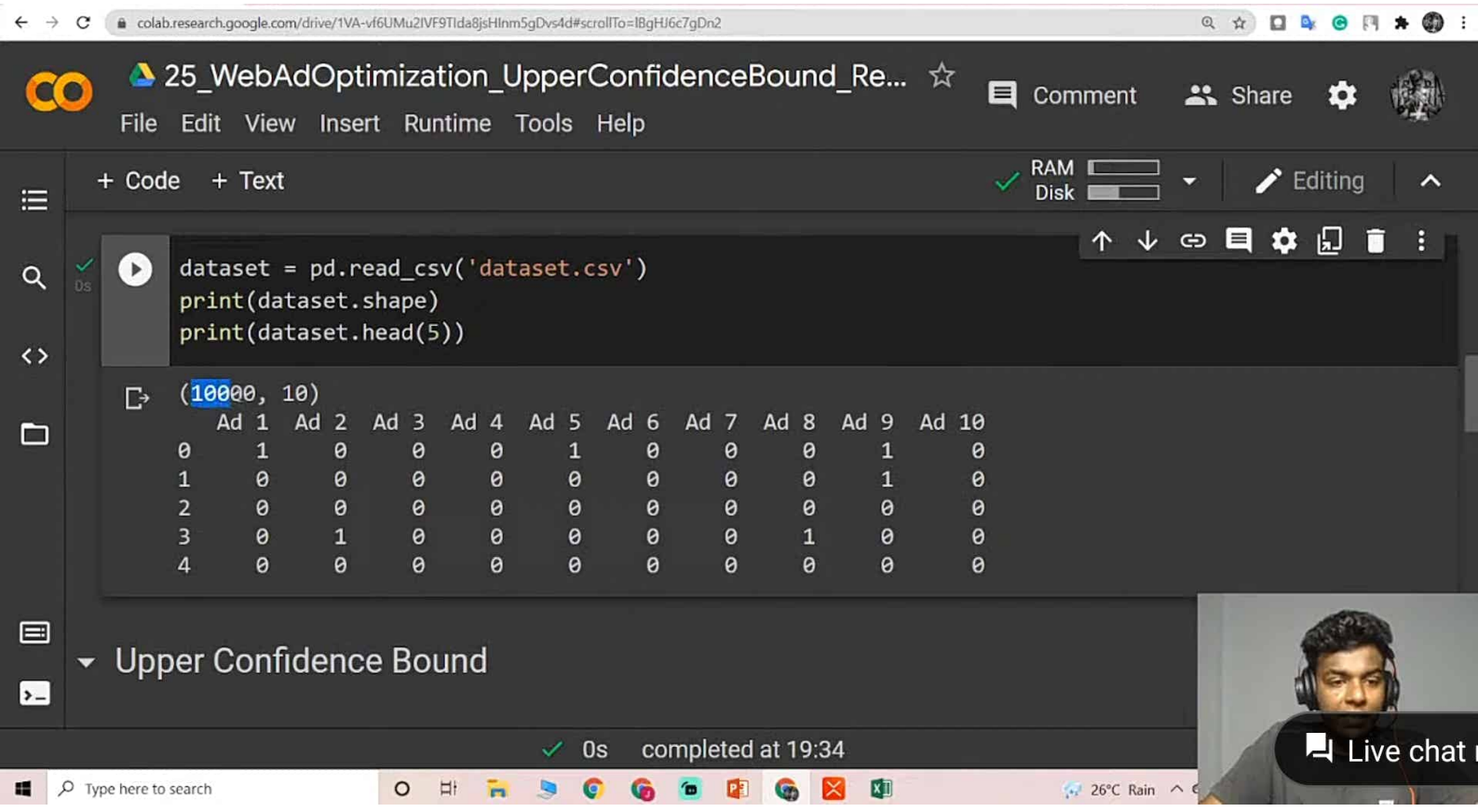

From Multiple Ads.
1
Finding the Problem - Application
Finding the Click Through Rate - CTR of certain Ad.
Input: Different Ads. with its CTR
2
Collecting Dataset
Web Ad. Dataset
3
Load and Summarize Dataset
Problem: Multi-Armed Bandit
1
You gonna select your department for your Higher studies in 35 Days
2
There are 5 Departments
D1
Medical
D2
Electronics/Electrical Engineering
D3
Computer Science Engineering
D4
Physics
D5
Chemical
Story - Example
3
You don't know how you will be happy and

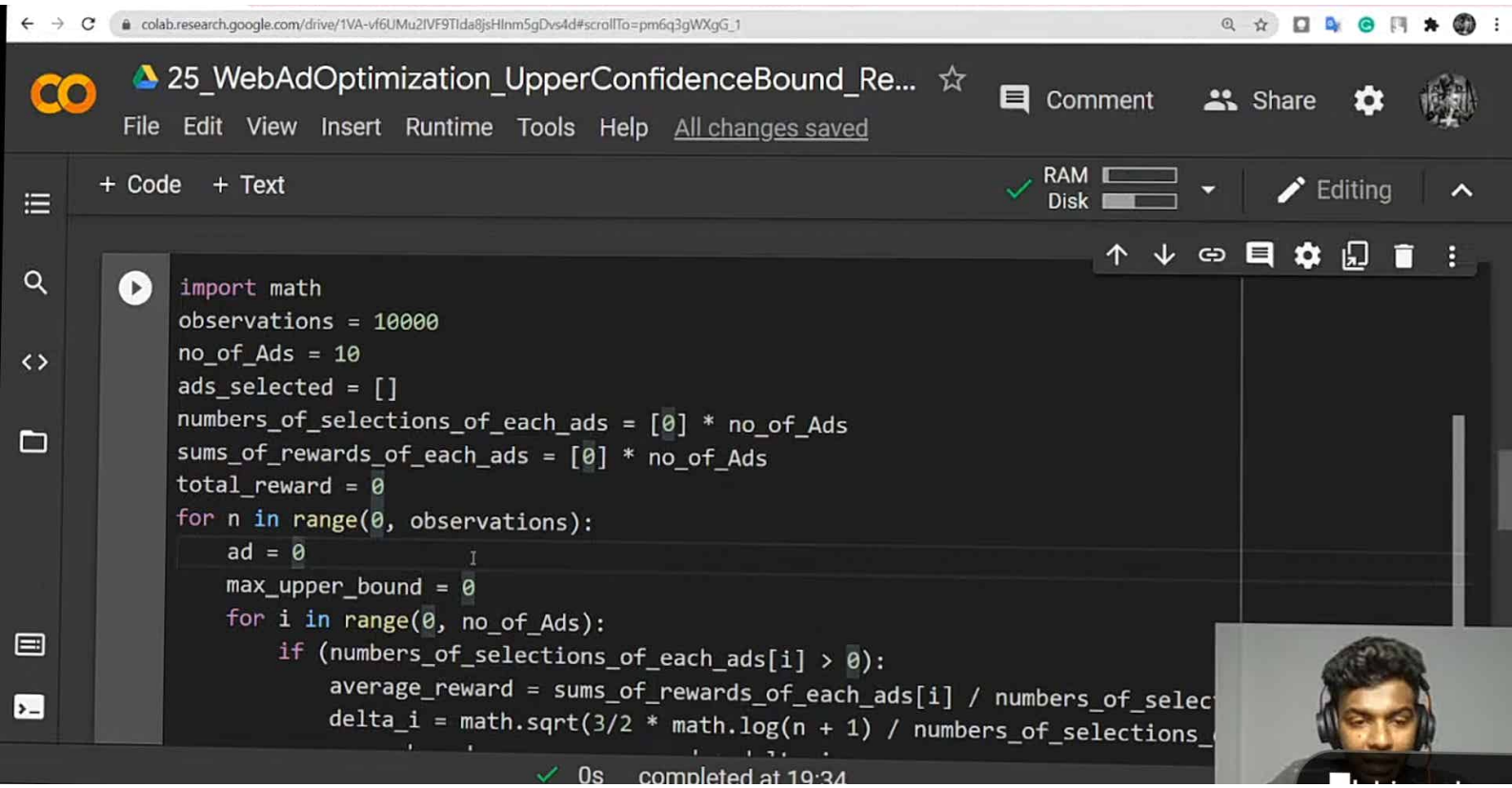

D4
Physics
D5
Chemical
Story - Example
3
You don't know how you will be happy and job opportunity in 5 Departments
4
Goal to pursue department from which department you will be more happy and with good job opportunity
we are going to devise a sequence of experiencing the happiness in a way to maximize the yield or rewards
5
Concept
Exploration
Exploring the curriculum of each and every course to find out which is giving you happy to study and which you think there is more opportunity for the job
Exploitation
Once the department is chosen, its time to pursuing the department in order to feel the happy and to get job in that department
4
Algorithm
1. Definition
It is a deterministic algorithm - there is no factor of uncertainty or probability
Reinforcement Learning that focuses on exploration and exploitation based on a confidence boundary that the algorithm assigns to each department on each round of exploration
1
Each department is assumed to have a uniform Confidence Interval and a success distribution
This Confidence Interval is a margin of success rate distributions which is the most certain to consist of the actual success rate distribution of each department which we are unaware of in the beginning
A department is randomly chosen to Study,

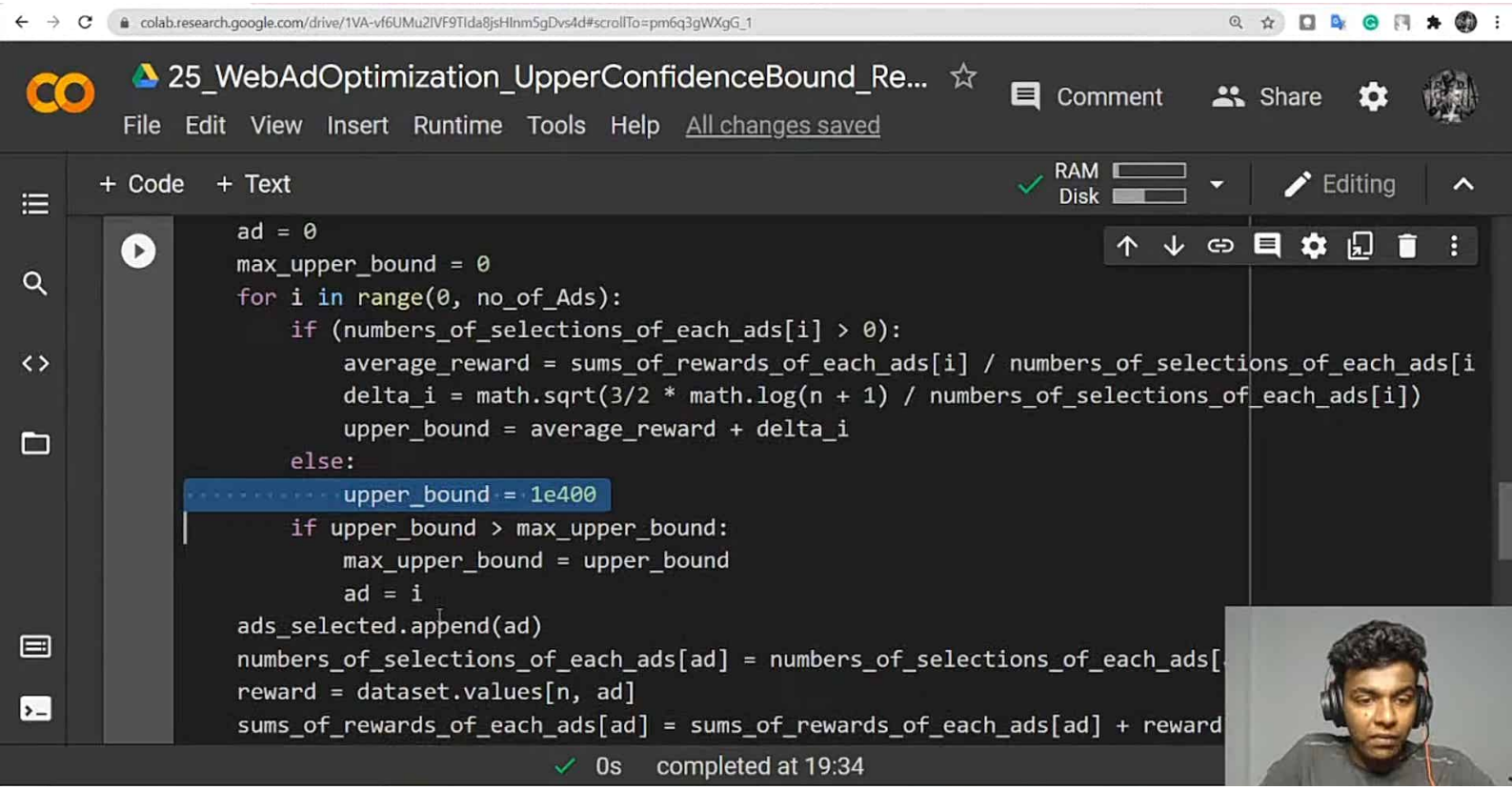
- Upper Confidence Bound
  - 2. Steps
    - 2: as initially, they have all the same confidence intervals
    - 3: Based on whether the department gave a reward or not, the Confidence Interval shifts either towards or away from the actual success distribution and the also converges or shrinks as it has been explored thus resulting in the Upper bound value of the confidence Interval to also be reduced
    - 4: Based on the current Upper Confidence bounds of each of the department, the one with the highest is chosen to explore in the next round
    - 5: Steps 3 and 4 are continued until there are sufficient observations to determine the upper confidence bound of each department. The one with the highest upper confidence bound is the department with the highest success rate
  - 3. Math Steps
    - 1: Two values are considered for each round of exploration of a department
      - The number of times each department has been selected till round n
      - The sum of rewards collected by each department till round n
    - 2: At each round, we compute the average reward and the confidence interval of the department i up to n rounds as follows
      - $\bar{r}_i(n) = \frac{R_i(n)}{N_i(n)}$ Average reward
      - $[\bar{r}_i(n) - \Delta_i(n), \bar{r}_i(n) + \Delta_i(n)]$

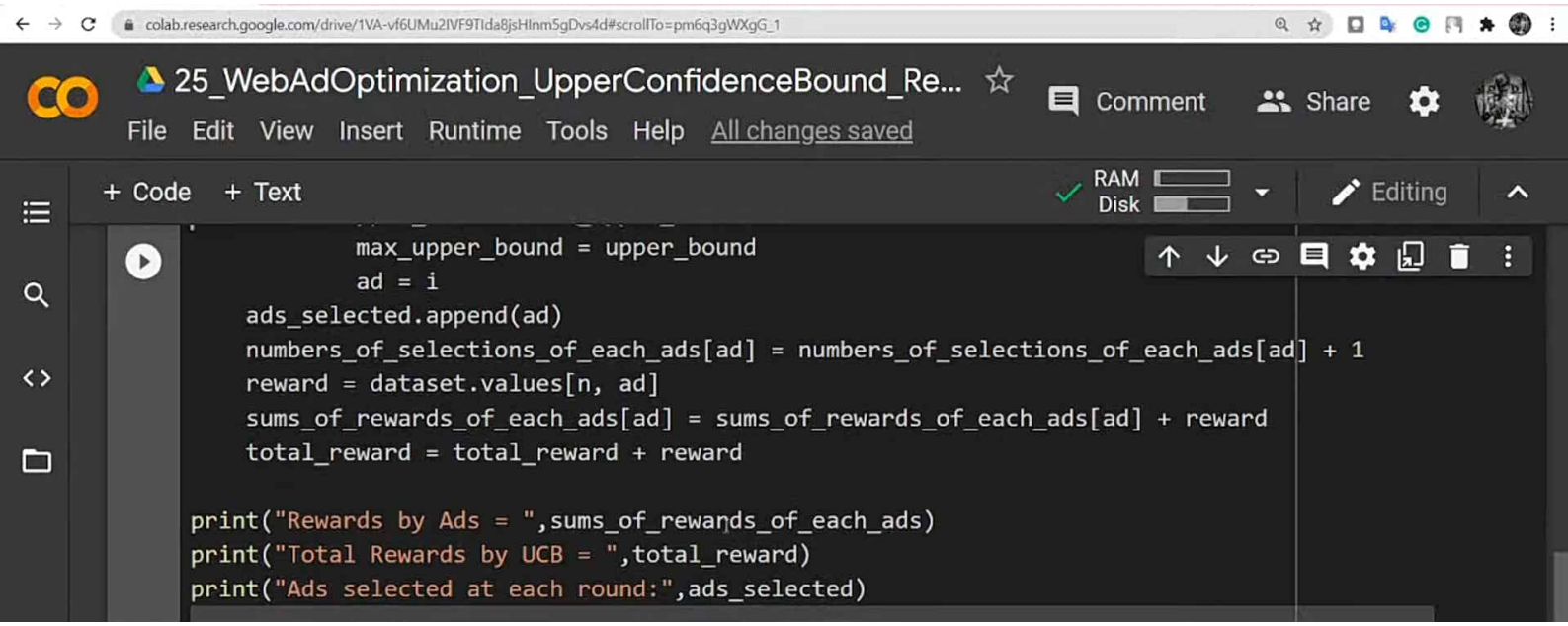

$N_i(n)$
Average reward
3. Math Steps
2
At each round, we compute the average reward and the confidence interval of the department i up to n rounds as follows
$[\bar{r}_i(n) - \Delta_i(n), \bar{r}_i(n) + \Delta_i(n)]$
$\Delta_i(n) = \sqrt{\frac{3}{2}\frac{\log(n)}{N_i(n)}}$
Confidence Interval
3
The department with the maximum UCB is selected
$\bar{r}_i(n) + \Delta_i(n)$
the maximum UCB
5
Visualizing result


colab.research.google.com/drive/1VA-vf6UMu2lVF9TIda8jsHInm5gDvs4d#scrollTo=y98nA5UdU6Hf
25_WebAdOptimization_UpperConfidenceBound_Re...
Comment
Share
File Edit View Insert Runtime Tools Help
+ Code + Text
RAM
Disk
Editing
Importing the basic libraries
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
Load Dataset from Local Directory
from google.colab import files
uploaded = files.upload()
Importing the dataset
0s completed at 19:33
Live chat


colab.research.google.com/drive/1VA-vf6UMu2IVF9TIda8jsHInm5gDvs4d#scrollTo=lBgHJ6c7gDn2
25_WebAdOptimization_UpperConfidenceBound_Re...
Comment
Share
File Edit View Insert Runtime Tools Help
+ Code + Text
RAM
Disk
Editing
dataset = pd.read_csv('dataset.csv')
print(dataset.shape)
print(dataset.head(5))
(10000, 10)
Ad 1 Ad 2 Ad 3 Ad 4 Ad 5 Ad 6 Ad 7 Ad 8 Ad 9 Ad 10
0 1 0 0 0 1 0 0 0 1 0
1 0 0 0 0 0 0 0 0 1 0
2 0 0 0 0 0 0 0 0 0 0
3 0 1 0 0 0 0 0 1 0 0
4 0 0 0 0 0 0 0 0 0 0
Upper Confidence Bound
0s completed at 19:34
Live chat
Type here to search
26°C Rain


colab.research.google.com/drive/1VA-vf6UMu2IVF9Tlda8jsHlnm5gDvs4d#scrollTo=pm6q3gWXgG_1
25_WebAdOptimization_UpperConfidenceBound_Re...
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
Editing
import math
observations = 10000
no_of_Ads = 10
ads_selected = []
numbers_of_selections_of_each_ads = [0] * no_of_Ads
sums_of_rewards_of_each_ads = [0] * no_of_Ads
total_reward = 0
for n in range(0, observations):
ad = 0
max_upper_bound = 0
for i in range(0, no_of_Ads):
if (numbers_of_selections_of_each_ads[i] > 0):
average_reward = sums_of_rewards_of_each_ads[i] / numbers_of_selec
delta_i = math.sqrt(3/2 * math.log(n + 1) / numbers_of_selections_
0s completed at 19:34


colab.research.google.com/drive/1VA-vf6UMu2IVF9TIda8jsHInm5gDvs4d#scrollTo=pm6q3gWXgG_1
25_WebAdOptimization_UpperConfidenceBound_Re...
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
Editing
ad = 0
max_upper_bound = 0
for i in range(0, no_of_Ads):
if (numbers_of_selections_of_each_ads[i] > 0):
average_reward = sums_of_rewards_of_each_ads[i] / numbers_of_selections_of_each_ads[i
delta_i = math.sqrt(3/2 * math.log(n + 1) / numbers_of_selections_of_each_ads[i])
upper_bound = average_reward + delta_i
else:
upper_bound = 1e400
if upper_bound > max_upper_bound:
max_upper_bound = upper_bound
ad = i
ads_selected.append(ad)
numbers_of_selections_of_each_ads[ad] = numbers_of_selections_of_each_ads[
reward = dataset.values[n, ad]
sums_of_rewards_of_each_ads[ad] = sums_of_rewards_of_each_ads[ad] + reward
0s completed at 19:34


colab.research.google.com/drive/1VA-vf6UMu2IVF9TIda8jsHInm5gDvs4d#scrollTo=pm6q3gWXgG_1
25_WebAdOptimization_UpperConfidenceBound_Re...
Comment
Share
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
RAM
Disk
Editing
max_upper_bound = upper_bound
ad = i
ads_selected.append(ad)
numbers_of_selections_of_each_ads[ad] = numbers_of_selections_of_each_ads[ad] + 1
reward = dataset.values[n, ad]
sums_of_rewards_of_each_ads[ad] = sums_of_rewards_of_each_ads[ad] + reward
total_reward = total_reward + reward
print("Rewards by Ads = ",sums_of_rewards_of_each_ads)
print("Total Rewards by UCB = ",total_reward)
print("Ads selected at each round:",ads_selected)

colab.research.google.com/drive/1VA-vf6UMu2lVF9TIda8jsHInm5gDvs4d#scrollTo=wOuthXLJT0GI

25_WebAdOptimization_UpperConfidenceBound_Re...

File Edit View Insert Runtime Tools Help Saving...

Comment Share

+ Code + Text

RAM Disk

Editing

```
plt.hist(ads_selected)
plt.title('Histogram of ads selections')
plt.xlabel('Ads')
plt.ylabel('Number of times each ad was selected')
plt.show()
```