Total Questions: 3 + 1
Total Marks: 350

All the files required for submissions must be named as described in the question and zipped. Submit each individual zip file for the question and upload it in the google form.

Q1. Adventures of Indiana Jones and his OR analyst: (100 marks)

Indiana Jones has hired you to assist him in his adventures as an unpaid intern. You both raid the tomb of an ancient egyptian ruler and find a huge treasure of idols, weapons and other artifacts. You see an opportunity in this to get some returns out of this dangerous internship by selling some of these. But, this evil thought is detected by the Goddess of the tomb and she starts attacking you both with snakes and scorpions. You quickly start thinking of ways to get the maximum value out of the available treasure. Coincidently (or conveniently!), you have access to an egyptian-era highly accurate weighing machine (thanks to those aliens!). You quickly weigh all the treasure and estimate the value of each artifact while Indiana Jones is fighting the enemies. You wish to take all the treasure but your little bag will break if you take too many of the artifacts. Your job is to decide which artifacts to take with you without breaking the bag while maximizing the value.

Data –

The algorithm you will be building will be tested in various scenarios. Each scenario consists of a list of items with their values and their weights. These parameters along with the number of items and the capacity of the bag change for every scenario. The first line will contain the number of artifacts in the tomb (n) and the maximum weight limit of your backpack (k). The n lines after the first line will have 2 values each. First element is the value of the artifact (v) and the next is the weight (w), as shown below.

```
n k

v_1 w_1

v_2 w_2

...

v_n w_n
```

Download the data from the link - https://drive.google.com/file/d/1x5ClyX_sGYPtRFuxt2FNlhuWnUT2CtGg/view?usp=sharing

Submission –

1. You will be provided with 13 (Case1-Case13) different scenarios. Each scenario will have a file as an input which is described as above. You can write a LP model or an algorithm in python to **solve cases Case4-Case12** and maximize the value in the bag without breaking the bag. Optimal solutions for Case1,Case2,Case3 are given in Table 1 so that you can test your algorithm. Also see SampleSubmission file in the folder for more details.

2. The python code/model you write must output a string of 1's and 0's (separated by a comma) and the value in the bag. 1 represents that an item is chosen while 0 depicts that it is not. For instance - consider a scenario with 5 objects. If we get a string 0,1,1,0,1, it means that objects 2,3 and 5 are chosen. The numbering of the objects is as per the data file. First object is 1 and so on. You will have to compile these strings and the value obtained in an excel file in following format (Table 1) -

Case	Solution String	Value of the bag
Case1	0,0,1,0,0,1,0,1,0,0,0, 0,1,1,0,0,0,0,0	12248.0
Case2	0,0,1,0,1,0,1,0,1,0,1, 0,1,0,0,0,1,0,1,0,1,0, 0,0,0,0,0,0,0,0	99798.0
Case3	0,0,1,0,1,0,0,0,1,0,0, 0,0,0,0,0,0,0,0,0,0,0, 0,0,0,0,1,0,1,0,0,0,1, 0,0,0,1,0,1,0	99924.0
...	...	...

Table 1

3. Name the excel file as TeamName_Q1

4. Sometimes LPs may take a long time to give an optimal solution. So, some alternate methods can be used to get a “good” solution in a short time which may not be optimal. For Case 13, can you generate a simple heuristic method to get a good solution? Please explain the method in short and implement it in python. Also provide the solution obtained in the same file as above. Provide a short explanation and copy-paste the code in a pdf file named TeamName_Q1.pdf.

Participants will be provided a google form to submit these responses.

Please note that we will need the list of 1's and 0's and also the optimal values. The responses will be checked using a python script If the solution string does not give the value of the bag you reported or if it requires more capacity than the bag, that particular scenario will be given 0 marks.

Evaluation –

1. Each scenario is worth 10 marks. Your total marks for this question will be the sum of all the marks for the scenarios (maximum 100). Marking scheme for an individual scenario is described below.
2. We (organizers of GC) have optimal solutions for cases 4-12. The following scheme gives marking rules for an individual scenario. If your solution is feasible and the value reported by you corresponds to that solution then you will be given marks as follows:
    a. If your value = 50-70% of optimal solution - 3 marks
    b. If your value = 71-99% of optimal solution - 7 marks
    c. If your value = 99-100% of optimal solution - 10 marks
3. For Case13 the maximum value out of all the submissions will be taken as the optimal solution and the grading will be as shown above.

Q2. Machine Learning (ML): (100 marks)

You were hired as a data-analyst at a chemical plant. They wish to model their plant's output based on the inputs which they can control. The inputs comprise of the Temperature, flow rate and 3 different concentrations. The outputs are the product quality and the cost. You are supposed to use ML techniques to model 2 the outputs. (No chemical engineering knowledge is necessary! Just ML!)

Data -

You will be given 2 data sets. train.csv and test.csv. The train.csv file will have inputs and the outputs while the test.csv will have only the inputs. You need to build a model using data in train.csv and predict the outputs for inputs in test.csv. You have to submit this csv in a google form which will be floated.

Download the data from the link below - https://drive.google.com/file/d/1oSZ9fi8eBBbfASB3F1qZFSjifS_rvmcl/view?usp=sharing

Submission -

The participants are expected to write their prediction in a csv file named TeamName_Q2.csv It must contain 2 columns. The first column must be of Quality and second of Cost in the following format. Please see sample submission for more details.

Quality	Cost
0.00	0.00
...	...

Note that the prediction for each row in test.csv must correspond to the same row in TeamName_Q2.csv (your submission). Otherwise, that will affect your rmse value severely. Zip the file.
You have to predict both the quality and the cost. Each prediction (Quality and Cost) will be judged separately and will have an equal weightage of maximum 50 marks.

Evaluation -

We will compare the predictions submitted by the participants with the ones which we have for test.csv. The submissions will be ranked based on RMSE (Root Mean Squared Error) of the predictions. Lesser the RMSE, higher on the ranking list is the submission. The submission with least RMSE will be placed first.

The score for one output will be
$(max_{rmse} - your\ rmse)/(max_{rmse} - min_{rmse}) \times 50$.

Qn 3a. Marks: 50

This Question Needs to be submitted before 3 PM.

You are given a fair coin, come up with a strategy to generate an integer between 1 and 10 (both included) with equal probability.

You cannot use any standard function to generate random numbers, except only one (use it any number of times) to generate the output of coin toss, i.e 0 or 1.

Submission: 3.py file

This file should contain a function which returns a value between 1 and 10.

Evaluation: Your function will be called a large number of times (greater than 1000) and frequency of output will be calculated.

Marks= 50 - 500*(standard deviation of probabilities of [1,10])

Qn 3b. Marks: 30

Three friends Akash, Kiran, and Tushar try the above problem and come up with strategies as described below. Say whether the strategies provide valid output and if they are correct.

Akash: Toss the coin 10 times, return the sum.

Kiran: Let S be the set of integers from 1 to 10. Toss the coin for each integer once, if heads, do nothing, if tails eliminate it from S. Repeat until one integer remains. Return this integer.

Tushar:
Take a thread of length 15 cm and two match sticks.
Lay the thread straight. Put one Match stick at the start and other at the end.
Toss a coin.
If heads, shift the match stick at start towards end stick by half the distance between start stick and end stick.
If tails, move end stick towards start stick by half the distance between start stick and end stick.
Repeat till the distance between start and end stick is less than 1 cm.
Output (start+end)/2

Example:
Position of Start match stick, position of end Match Stick
Initial Position --->0,15
Tails --->7.5, 15.0
Head ---->7.5,11.25

Tails---->7.5, 9.375
Head ----->8.4375, 9.375

Now the difference is less than one, hence output=6

Submission: Explanation with Proof.
Evaluation: 10 for each case if answer and Proof is correct.

Question No. 3c: Marks 10
In which scenario will the strategy of Tushar be correct?
Submission: Explanation with Proof.
Evaluation: 10 if answer and Proof is correct.

Question No. 3d: Marks 10
Does taking thread of length “15” bear any significance? Or should it be “16” cm?
Submission: Explanation with Proof.
Evaluation: 10 if answer and Proof is correct.

Q.4. Bonus Question : 50 Marks

Suraj and Chand, each have a pile of **n** chips. (n >10)

Suraj rolls a die, if an even number appears, he eats two chips from the pile, if odd, he transfers one chip from his pile to that of Chand. Then Chand does the same, if die face on top is even, he eats from his pile, else transfers. If we say that this game is disadvantageous for the second player, do you agree?

A change to this game was suggested as:

Suraj rolls a die, if an even number appears, he eats two chips from the pile, if odd, he eats 1. Then Chand does the same. The first player to remove all the chips from their own pile is the winner. Is the game fair now?

If a player is left with a single chip and an even number appears, it's a pass.

Submission: Explanation with Proof.

Evaluation: Correct Answer and Proof for each case will be awarded 25 marks.