**Python:**

Total Marks: 20  
Each question 10 marks

**Question: 1**

1. Write an efficient algorithm that searches for a value target in an m x n integer matrix. This matrix has the following properties:
   1. Integers in each row are sorted from right to left.
   2. The first integer of each row is greater than the last integer of the previous row.

**Example**-:

                                        Input: matrix = [[1,3,5,7],[10,11,16,20],[23,30,34,60]], target = 3

                                         Output: True

**Question: 2**

**2.** Write a program that takes a string as input, and counts the frequency of each word in the string, there might  be repeated characters in the string. Your task is to find the highest frequency and returns the length of the  highest-frequency word.

**Note -** You have to write at least 2 additional test cases in which your program will run successfully and provide  an explanation for the same.

Example input - string = “write write write all the number from from from 1 to 100”

Example output - 5

Explanation - From the given string we can note that the most frequent words are “write” and “from” and  the maximum value of both the values is “write” and its corresponding length is 5

**Machine learning:**

Total Marks: 30  
Each question 15 marks

**Question: 1**

1. Imagine you have a dataset where you have different Instagram features like u **sername , Caption , Hashtag , Followers , Time\_Since\_posted , and likes ,** now your task is to predict the number of likes and Time Since posted and the rest of the features are your input features. Now you have to build a model which can predict the number of likes and Time Since posted.

[Dataset](https://www.kaggle.com/datasets/rxsraghavagrawal/instagram-reach) This is the Dataset You can use this dataset for this question.

**Question: 2**

2.

1. Explain how you can implement ML in a real world application.
2. Train an SVM regressor on : [Bengaluru housing dataset](https://www.kaggle.com/datasets/amitabhajoy/bengaluru-house-price-data)

   Must include in details:

  - EDA

  - Feature engineering

**Deep Learning :**

Total Marks: 40  
Each question 20 marks

**Question: 1**

1. Train a Pure CNN with less than 10000 trainable parameters using the MNIST Dataset having minimum validation accuracy of 99.40%

**Note -**Code comments should be given for proper code understanding.

**Question: 2**

2.

1. Explain how you can implement DL in a real-world application.
2. Train an industry safety Detection model that will detect the helmet wearing by the employee: [Data link](https://www.kaggle.com/datasets/trainingdatapro/helmet-detection?select=img)

* **Note:** You are free to choose frameworks as per your understanding

**Natural Language Processing:**

Total Marks: 40

Each question 20 marks

**Question: 1**

1. **Data Set Link: (Take Any Dataset but avoid using benchmark data)**

**Perform Below Listed Preprocessing Task in proper order as per your dataset using torch or tensorflow.**

1. Emoji Removal
2. Lemmatization
3. Stemming
4. Word Tokenization
5. Grammar Correction
6. Http Links Removal
7. Stop Words Removal
8. Sentence Tokenization
9. Lower casing
10. Remove white spaces
11. Text Normalization
12. Part of speech tagging

**Question: 2**

2. Dataset Link: (Take Any Dataset but avoid using benchmark data)

**Tech** : Pytorch or Tensorflow

Take any Dataset and perform any one task. (You are free to choose Tasks as per your understanding)

* Tasks: Sentiment Analysis , Text Classification , Text Generation, Machine Translation, Text Summarization , Question Answering.