ANSWERS FOR THE QUESTIONS GIVEN IN ASSIGNMENT

Q1. Could we sum it?

Let Ankur has given a large weight W, and a list of smaller weights in an array. He needs to write a code in order to find "can we form weight W or not, using smaller weights". He only knows dp solution. Could you write a code solution for him without using dp.

A1. The python code for the same is given below:

```
def check_weight_formability(W,arr):
    if W == 0:
        return True
    if arr == []:
        return False
    if W == arr[0]:
        return True
    if W > arr[0]:
        W = W - arr[0]
        arr.pop(0)
        return check_weight_formability(W, arr)
    arr.pop(0)
    return check_weight_formability(W, arr)

W = int(input("Please enter the large weight: "))
given_list = []
num_ele = int(input("Please enter the desired number of smaller weights (less than
                    12): "))
if num_ele <= 12:
    print("Please enter the smaller weights: ")
    for i in range(0, num_ele):
        ele = int(input())
        given_list.append(ele)
    given_list.sort(reverse = True)
    print(check_weight_formability(W,given_list))
else:
     print("You have exceeded maximum number of inputs!!!")
```

Q2. How Do You Handle Missing or Corrupted Data in a Dataset?

A2. Missing or corrupt data must be handled to ensure that the values we obtain from the analysis make sense. There are multiple ways to handle missing/corrupted data using python. One can remove the entire row or column which contains such values in them. However, this can prove costly as the dataset size may considerably decrease doing so resulting in very low sample size available for analysis. Another way to handle such an issue is using imputations. Imputation is the process by which we can replace the missing values like “NaN” or “Null” by some other value (for e.g., the mean, median, or mode). Whatever method we choose, it is mandatory to understand the consequences of the

method we choose. For e.g., when we do imputation, it may be possible that we used mean to replace the missing values, but replacing using a mode value or regression value was a batter option.

Q3. What Are the Three Stages of Building a Model in Machine Learning?

A3. The three stages of building a model in Machine Learning are:

1. Data preprocessing : This step involves the data cleaning, dealing with missing values or corrupted values, feature selection, creating dummy variables for categorical data, normalisation of data, and splitting up the data to train, validation and test sets.
2. Training the model: This step involves identifying a suitable algorithm and using it on the training dataset to train the model using specific hyperparameters associated with the corresponding model.
3. Testing and validation: The trained model from the above step is used to test in validation dataset in real-time to ensure that the model is not overfitting. This is an essential step as overfitting can lead to increased accuracy but it loses the sense of the real-world problem in hand. If the results are not satisfactory, the hyperparameters need readjustment and this marks the beginning of an iterative procedure in solving the problem. After it is ensured that the model has not performed overfitting and that the model gives an accuracy which is satisfactory, the model is to be applied on test dataset to understand the accuracy of the model in real-world data.