DECISION TREE –CLASSIFICATION

Problem Statement or Requirement:

A client's requirement is, he wants to predict the purchased/not purchased based on the several parameters. The Client has provided the dataset in csv file.

As a data scientist, you must develop a model which will predict the purchased/not purchased.

Identify your problem statement:

Stage 1:

They provide dataset in csv file. So we shall take machine learning.

Stage 2:

requirement is clear. Input and output are present here. So we shall take Supervised learning.

Stage 3:

Then out put's are categorical value so we take classification.

2 X 2 MATRIX

NOT PURCHASED (0) - 257
PURCHASED - 143
TOTAL - 400

CONFUSION MATRIX TABLE

71	8
3	38

```
print(clf_report)
```

	precision	recall	f1-score	support
0	0.96	0.90	0.93	79
1	0.83	0.93	0.87	41
accuracy			0.91	120
macro avg	0.89	0.91	0.90	120
weighted avg	0.91	0.91	0.91	120

6 Types of Evaluation Matrix – Interview Questions

1. Accuracy

- ***What is the percentage of correct classification of both purch ased and not purchsed to the total input of the test set ?***
- ***Over all performance of the model ?***
- ***What is the accuracy of the classification problem statement ?***
- ***What is the overall performance of the model of Decision Tre e ?***

ANSWERS: 0.91

2. Recall

- *What is the percentage of correct classification of not purch ased to the*
  total input of not purchased in the test set ?

- *What is the correct classification of not purchased ?*

ANSWERS: 0.90

- *What is the percentage of correct classification of purchase d to the total*
  in put of purchased in the test set ?

- *What is the correct classification of purchased ?*

ANSWERS: 0.93

3. ***Precision***

- *What is the percentage of correct classification of (not purc hased) to sum of correctly classified as (not purchased) in the test set ?*

ANSWERS: 0.96

- *What is the percentage of correct classification of (purchase d) to sum of*
  correctly classified as (purchased) in the test set ?

ANSWERS: 0.83

4. ***F1- Score***

- *What is the overall performance of not purchased ?*

ANSWERS: 0.93

- ***What is the overall performance of purchased ?***

ANSWERS: 0.87

5. ***Macro Average***

- ***What is the average performance of precision (correctly and wrongly)***
  classified ?

ANSWERS: 0.89

- ***What is the average performance of Recall (correctly and wro ngly)***
  classified ?

ANSWERS: 0.91

- ***What is the average performance of F1-Meassure (correctly and***
  wrongly) classified ?

ANSWERS: 0.90

6. ***Weighted Average***

- ***What is the sum of product of proportion rate (weight of eac h***
  class-precision) ?

ANSWERS: 0.91

- ***What is the sum of product of proportion rate (weight of eac h***

  class-Recall) ?

ANSWERS: 0.91

- ***What is the sum of product of proportion rate (weight of eac h***

  class-F1-Meassure) ?

ANSWERS: 0.91