Scenario 1:

You have trained a binary classification model and used it to make predictions on a test set of 1000 observations. The test set contains 400 positive examples and 600 negative examples. The model's confusion matrix is as follows:

|  | **Predicted Positive** | **Predicted Negative** |
| --- | --- | --- |
| Actual Positive | 350 | 50 |
| Actual Negative | 100 | 450 |

1. What is the model's accuracy on this test set?

Accuracy=TP+TN/FP+TP+TN+FN=350+450/50+450+350+100=0.8421

1. What is the model's precision for positive examples?

PRESICION=TP/TP+FP=350/350+100=0.7777

1. What is the model's recall for positive examples?

RECALL=TP/TP+FN=350/350+50=0.875

1. What is the model's false positive rate?

FPR=FP/TN+FP=100/450+100=0.1818

1. What is the model's F1-score for positive examples?

F1 SCORE=2\*PRECISION \*RECALL/PRECISION+RECALL=2\*0.77\*0.87/0.77+0.87=.8169

Scenario 2:

You have trained a multi-class classification model and used it to make predictions on a test set of 1000 observations. The test set contains 100 observations for each of the classes A, B, C and D. The model's confusion matrix is as follows:

|  | **Predicted A** | **Predicted B** | **Predicted C** | **Predicted D** |
| --- | --- | --- | --- | --- |
| Actual A | 80=TPA | 5=FP | 10 | 5 |
| Actual B | 5 | 90 | 2 | 3 |
| Actual C | 8 | 3 | 80 | 9 |
| Actual D | 7 | 2 | 5 | 86 |

1. What is the model's overall accuracy on this test set?

Overall accuracy=0.84

1. What is the model's precision for class A?

Model precision for class A=0.8

1. What is the model's recall for class D?

Model recall for class d=0.86

1. What is the model's F1-score for class C?

F1 score for class c =2\*precision \*recall of c /precision +recall of c=2\*0.82\*0.8/0.82+0.8=0.8098

1. What is the model's macro-average recall across all classes?

macro-average recall=0.8+0.9+0.8+0.86/4=0.84

In both of the above scenarios the true positive(TP), False Positive(FP), True Negative(TN), False Negative(FN) can be calculated and used to derive the respective metrics like accuracy, precision, recall and F1-score.