

## MANUAL CALCULATION FOR THE CLASSIFICATION REPORT :

Random forest algorithm:

Given confusion matrix:

```
[[72  7]
 [ 3 38]]
```

Total counts in the test set=120

True non-purchased=72

False non-purchased =7

True purchased=38

False purchased=3

Classification report :

	precision	recall	f1-score	support
0	0.96	0.91	0.94	79
1	0.84	0.93	0.88	41
accuracy			0.92	120
macro avg	0.90	0.92	0.91	120
weighted avg	0.92	0.92	0.92	120

1.Accuracy:

$$\begin{aligned} &= \frac{T(\text{purchased})+T(\text{non-purchased})}{\text{sum of all the entries in the given matrix}} \\ &= \frac{72+38}{72+38+3+7} \\ &= \frac{110}{120} = 0.92 \end{aligned}$$

2. Recall of non-purchased:

$$\begin{aligned} &= \frac{T(\text{non-purchased})}{T(\text{non-purchased}) + F(\text{non-purchased})} \\ &= \frac{72}{72+7} = 0.91 \end{aligned}$$

3. Recall of purchased:

$$\begin{aligned} &= \frac{T(\text{purchased})}{T(\text{purchased})+F(\text{purchased})} \\ &= \frac{38}{38+3} = \frac{38}{41} = 0.93 \end{aligned}$$

4. Precision of non-purchased:

$$\begin{aligned} &= \frac{T(\text{non-purchased})}{T(\text{non-purchased})+F(\text{purchased})} \\ &= \frac{72}{72+3} = 0.96 \end{aligned}$$

5. Precision of purchased:

$$\begin{aligned} &= \frac{T(\text{purchased})}{T(\text{purchased})+F(\text{non-purchased})} \\ &= \frac{38}{38+7} = 0.84 \end{aligned}$$

6. F1-score of non-purchased:

$$\begin{aligned} &= \frac{2 * \text{recall of non-purchased} * \text{precision of non-purchased}}{\text{recall of non-purchased} + \text{precision of non-purchased}} \\ &= \frac{2 * 0.91 * 0.96}{0.91 + 0.96} = \frac{1.7472}{1.87} \\ &= 0.94 \end{aligned}$$

7. F1-score of purchased:

$$\begin{aligned} &= \frac{2 * \text{recall of purchased} * \text{precision of purchased}}{\text{recall of purchased} + \text{precision of purchased}} \\ &= \frac{2 * 0.93 * 0.84}{0.93 + 0.84} = \frac{1.5624}{1.77} \\ &= 0.88 \end{aligned}$$

8. Macro average of precision:

$$\begin{aligned} &= \frac{\text{precision}(\text{non-purchased}) + \text{precision}(\text{purchased})}{2} \\ &= \frac{0.96 + 0.84}{2} = 0.90 \end{aligned}$$

9. Macro avg of recall:

$$\begin{aligned} &= \frac{\text{recall}(\text{non-purchased}) + \text{recall}(\text{purchased})}{2} \\ &= \frac{0.91 + 0.93}{2} = 0.92 \end{aligned}$$

10. Macro avg of f1-score:

$$= \frac{f1(\text{non-purchased}) + f1(\text{purchased})}{2}$$

$$= \frac{0.94 + 0.88}{2} = 0.91$$

#### 11. Weighted average of precision:

$$= (\text{precision}(\text{non-purchased}) * (79/120)) + (\text{precision}(\text{purchased}) * (41/120))$$

$$= (0.96 * 0.66) + (0.84 * 0.34)$$

$$= 0.6336 + 0.2856$$

$$= 0.92$$

#### 12. Weighted avg of recall:

$$= [\text{recall}(\text{non-purchased}) * (79/120)] + [\text{recall}(\text{purchased}) * (41/120)]$$

$$= [0.91 * 0.66] + [0.93 * 0.34]$$

$$= 0.6006 + 0.3162$$

$$= 0.92$$

#### 13. Weighted avg of f1-score:

$$= [f1(\text{non-purchased}) * (79/120)] + [f1(\text{purchased}) * (41/120)]$$

$$= [0.94 * 0.66] + [0.88 * 0.34]$$

$$= 0.6204 + 0.2992$$

$$= 0.92$$

- We got the same values in both the classification report and manual check.
- In 10 classification algorithms , we got the best model in random forest algorithm (accuracy,precision,recall,f1-score values are high). So we proceed to next phase.