Classification report of random forest for the classification matrix array([[73, 4],

[6, 37]]

	precision	recall	f1-score	support
0	0.92	0.95	0.94	77
1	0.90	0.86	0.88	43
accuracy			0.92	120
macro avg	0.91	0.90	0.91	120
weighted avg	0.92	0.92	0.92	120

precision

1. Which gives us the percentage of correctly predicted positive class out of all positive classes – precision
   Lets say for class 0 the precision is .92 means that the model correctly predicted class 0 by a percent of 92percentage and remaing .08 percent event though the model predicted class 0 it was actually class 1.

   Same for class 1 .90 means means that the model correctly predicted class 1 by a percent of 90 percentage and remaing .10 percent event though the model predicted class 1 it was actually class 0.

   So if u break down the classification matrix array:

   Case 1 :where class 0 is treated as positive and class 1 is negative:

Predicted/ actual	Class 0(purchased).	Class 1(non purchased)
Class0	Tp=73	Fp=4
Class 1	Fn=6	tn=37

Hier precision = tp(count of coreectly predicted class 0)/(sum of tp+fp(all the positives predicted by the model)) =0.92

Predicted/ actual	Class 0(purchased).	Class 1(non purchased)
Class0	tn=73	Fn=4
Class 1	Fp=6	Tp=37

Case 2: where class 1 is treated as positive and class 0 is negative:

Hier precision = tp(count of coreectly predicted class 1)/(sum of tp+fp(all the positives predicted by the model)) =0.90.

recall

case 1 : when class 0 is positive

recall = total count of coreectly predicted true positives (class0) out of all actual positives(class0) =.95

case 2 : when class 1 is positive

recall = total count of coreectly predicted true positives (class1) out of all actual positives(class1) =.86

here recall is greater for class 0 than class 1 means the model correctly identified class 0 as positive than class 1 as positive.so that means the model will perform well in identifying class 0 than it identifying class 1.

F1 score

Which gives us a mean of precision and recall when either of them are altered to increase the model perforformance which results in decrease in recall or precision.for example

Lets say to increase the recall means we hv to make the model correctly predict more positive, while doing so it will also predict more false positive along with true positivesso when theres more false positives it will reduce precision.

Wen u want to increase precision u want to decrease the false positives,so model becomes extra careful to avoid false positives and in the process might miss true positives which leads to more false negatives and in turn recall decreases.

To avoid this , f1 score comes in play which calculates the harmonic mean btw 2 in an imbalanced dataframe.

Support

array([[73, 4],
 [6, 37]]

For the confusion matrix here lets say , the no of instances in class 0 is 73+4=77 the no of instances in class 1 is 6+37=43

so support gives the number/ count of actually occurring classes/instances .

For example if class 0 is purchased then 77 times something was actually purchased and 43 times something was actually “not purchased” (class 1)

accuracy

its the mean of total correctly predicted instances to the total no of instances

so its tp+tn(tp+tn+fp+fn)= (73+37)/120=.92

macro average

it calculate the mean of precision for both classes,mean of recall for both and mean of f1 for both. So it doesn’t show any bias btw 2 classes and is more suited for balanced data setes

macro precision =(precision 1 +precion 2)/ 2=.91

weighted average

its for imbalanced datasets

it gives importance to the support or the no. Of instances in each class

it gives greater importance to class with greater instance

so it might falsely predict the class with lower instances.