

Confusion Matrix

1. SVM - Classification

	(Not Purchased) Predicted Positive	(Not Purchased) Predicted Negative
(Not Purchased) Actual Positive	77 (TP)	2 (FP)
(Purchased) Actual Negative	23 (FN)	18 (TN)

1. What is the overall performance of the model prediction?

Accuracy Formula = $TP+TN/TP+TN+FP+FN$

$$= 77+18/ 77+2+23+18 =95/120$$

Accuracy → 0.79

2. Out of all the positive input (Purchased), how many are correctly predicted?

Recall Formula = $TN/TN+FN$

$$=18/18+23$$

$$=18/41$$

Recall → 0.44

3. Out of all the predicted positives (Purchased), how is actually positive

Precision Formula = $TN/TN+FP$

$$= 18/18+2$$

$$=18/20$$

Precision → 0.90

4. What is the overall performance of precision & recall?

Formula = $2*recall*precision/recall+precision$

$$= 2* 0.44*0.90/0.44+0.90$$

$$= 0.792/ 1.34$$

F1-score → 0.59

5. What is the average performance of recall, precision & F1 measure → Macro avg

Precision → 0.83, Recall → 0.71, F1-score → 0.73

Recall	Precision	F1-score
$=\text{recall}(\text{TN})+\text{recall}(\text{TP})/2$ $=0.44+0.97/2$ $=1.41/2$ $=0.71$	$=\text{precision}(\text{TP})+\text{Precision}(\text{TN})/2$ $=0.77+0.90/2$ $=1.67/2$ $=0.83$	$=F1(\text{TP})+F1(\text{TN})/2$ $=0.86+0.59/2$ $1.45/2$ $=0.73$

6. How many instances belong to each class → weighted avg

$\text{precision}(\text{TP}) * (\text{total count of TP} / \text{total input}) + \text{precision}(\text{TN}) * (\text{total count of TN} / \text{total input})$

Recall	Precision	F1-score
$=\text{recall}(\text{TP}) * (\text{total count of TP} / \text{total input}) / \text{recall}(\text{TN}) * (\text{total count of TN} / \text{total input})$ $=0.97 * (79/120) + 0.44 * (41/120)$ $=0.79$	$=\text{precision}(\text{TP}) * (\text{total count of TP} / \text{total input}) + \text{precision}(\text{TN}) * (\text{total count of TN} / \text{total input})$ $=0.77(79/120) + 0.90(41/120)$ $=0.81$	$=F1(\text{TP}) * (\text{total count of TP} / \text{total input}) + F1(\text{TN}) * (\text{total count of TN} / \text{total input})$ $=0.86(79/120) + 0.59(41/120)$ $=0.77$

Results:

Accuracy: 0.79

	Recall	Precision	F1-score	Support
0	0.97	0.77	0.86	79
1	0.44	0.90	0.59	41
macro avg	0.83	0.71	0.73	120
weighted avg	0.81	0.79	0.77	120

Random Forest

	(Not Purchased) Predicted Positive	(Not Purchased) Predicted Negative
(Not Purchased) Actual Positive	73 (TP)	6 (FP)
(Purchased) Actual Negative	5 (FN)	36 (TN)

1. What is the overall performance of the model?

$$\begin{aligned}\text{Accuracy Formula} &= \text{TP} + \text{TN} / \text{TP} + \text{TN} + \text{FP} + \text{FN} \\ &= 73 + 36 / 73 + 36 + 6 + 5 \\ &= 109 / 120 \\ &= 0.91\end{aligned}$$

2. Recall:

Out of all the positive inputs, how many does it predict correctly?

TP	TN
$\begin{aligned}&= \text{TP} / \text{TP} + \text{FP} \\ &= 73 / 73 + 6 \\ &= 73 / 79 \\ &= 0.92\end{aligned}$	$\begin{aligned}&= \text{TN} / \text{TN} + \text{FN} \\ &= 36 / 36 + 5 \\ &= 36 / 41 \\ &= 0.88\end{aligned}$

3. Precision:

Out of all the predicted positives, how many are actually positive?

TP	TN
$\begin{aligned}&= \text{TP} / \text{TP} + \text{FN} \\ &= 73 / 73 + 5 \\ &= 73 / 78 \\ &= 0.94\end{aligned}$	$\begin{aligned}&= \text{TN} / \text{TN} + \text{FP} \\ &= 36 / 36 + 6 \\ &= 36 / 42 \\ &= 0.86\end{aligned}$

4. F1 score: Overall performance of precision & recall

TP	TN
$2 * TP \text{ precision} * TP \text{ recall} / TP \text{ precision} + TP \text{ recall}$ $= 2 * 0.94 * 0.92 / 0.94 + 0.92$ $1.73 / 1.86$ $= 0.93$	$2 * TN \text{ precision} * TN \text{ recall} / TN \text{ precision} + TN \text{ recall}$ $= 2 * 0.86 * 0.88 / 0.86 + 0.88$ $= 1.51 / 1.74$ $= 0.87$

5. Macro Avg: What is the average of precision & recall

Recall	Precision	F1-score
$= TP \text{ recall} + TN \text{ recall} / 2$ $= 0.92 + 0.88 / 2$ $= 1.8 / 2$ $= 0.90$	$= TP \text{ precision} + TN \text{ precision} / 2$ $= 0.94 + 0.86 / 2$ $= 1.8 / 2$ $= 0.90$	$= TP \text{ f1 score} + TN \text{ f1 score} / 2$ $= 0.93 + 0.87 / 2$ $= 1.8 / 2$ $= 0.90$

6. Weighted Avg:

How many instances belong to each class

$TP(\text{recall}) * (\text{Total count of TP} / \text{total input}) + TN(\text{recall}) * (\text{Total count of TN} / \text{total input})$

Recall	Precision	F1-score
$= TP \text{ recall} (\text{Total TP} / \text{Total input}) + TN \text{ recall} (\text{Total TN} / \text{Total input})$ $= 0.92(79/120) + 0.88(41/120)$ $= 0.91$	$= TP \text{ precision} (\text{Total TP} / \text{total input}) + TN \text{ precision} (\text{Total TN} / \text{total input})$ $= 0.94(79/120) + 0.86(41/120)$ $= 0.91$	$= TP \text{ f1 score} (\text{Total TP} / \text{Total input}) + TN \text{ f1 score} (\text{Total TN} / \text{total input})$ $= 0.93(79/120) + 0.87(41/120)$ $= 0.91$

Results:

Accuracy: 0.91

	Recall	Precision	F1-score	Support
0	0.94	0.92	0.93	79
1	0.86	0.88	0.87	41
macro avg	0.90	0.90	0.90	120
weighted avg	0.91	0.91	0.91	120

3. Decision Tree

	(Not Purchased) Predicted Positive	(Not Purchased) Predicted Negative
(Not Purchased) Actual Positive	71 (TP)	8 (FP)
(Purchased) Actual Negative	3 (FN)	38 (TN)

1. Accuracy: What is the overall performance of the model

$$\begin{aligned}
 &= \text{TP} + \text{TN} / \text{TP} + \text{TN} + \text{FP} + \text{FN} \\
 &= 71 + 38 / 71 + 38 + 8 + 3 \\
 &= 109 / 120 \\
 &= 0.91
 \end{aligned}$$

2. Out of all the input, how many are correctly predicted

TP	TN
$ \begin{aligned} &= \text{TP} / \text{TP} + \text{FP} \\ &= 71 / 71 + 8 \\ &= 71 / 79 \\ &= 0.90 \end{aligned} $	$ \begin{aligned} &= \text{TN} / \text{TN} + \text{FN} \\ &= 38 / 38 + 3 \\ &= 38 / 41 \\ &= 0.93 \end{aligned} $

3. Out of all predicted values, how many does it predict correctly?

TP	TN
$ \begin{aligned} &= \text{TP} / \text{TP} + \text{FN} \\ &= 71 / 71 + 3 \\ &= 71 / 74 \\ &= 0.96 \end{aligned} $	$ \begin{aligned} &= \text{TN} / \text{TN} + \text{FP} \\ &= 38 / 38 + 8 \\ &= 38 / 46 \\ &= 0.83 \end{aligned} $

4. F1-Score: What is the overall performance of recall & precision?

TP	TN
$= 2 * \text{recall} * \text{precision} / (\text{recall} + \text{precision})$ $= 2 * 0.90 * 0.96 / (0.90 + 0.96)$ $= 1.73 / 1.86$ $= 0.93$	$= 2 * \text{recall(TN)} * \text{precision(TN)} / (\text{recall(TN)} + \text{precision(TN)})$ $= 2 * 0.83 * 0.93 / (0.83 + 0.93)$ $= 1.54 / 1.76$ $= 0.87$

Macro Avg:

Avg of precision & recall

Recall	Precision	F1-score
$= \text{Recall(TP)} + \text{recall (TN)} / 2$ $= 0.90 + 0.93 / 2$ $= 1.83 / 2$ $= 0.92$	$= \text{Precision (TP)} + \text{Precision (TN)} / 2$ $= 0.96 + 0.83 / 2$ $= 1.79 / 2$ $= 0.90$	$= \text{F1-score (TP)} + \text{F1-score (TN)} / 2$ $= 0.93 + 0.87 / 2$ $= 1.8 / 2$ $= 0.90$

weighted avg: How many instances belong to the each class

Recall	Precision	F1-score
$= \text{Recall(TP)} (\text{Total TP} / \text{Total input}) + \text{Recall (TN)} (\text{Total TN} / \text{total input})$ $= 0.90 * (79/120) + 0.93(41/120)$ $= 0.90 * 0.66 + 0.93 * 0.34$ $= 0.6 + 0.32$ $= 0.91$	$= \text{Precision (TP)} (\text{Total (TP)} / \text{total input}) + \text{Precision (TN)} (\text{Total (TN)} / \text{Total input})$ $= 0.96(79/120) + 0.83(41/120)$ $= 0.96 * 0.66 + 0.83 * 0.34$ $= 0.91$	$= \text{F1-score (TP)} (\text{Total TP} / \text{total input}) + \text{F1-score (TN)} (\text{Total TN} / \text{total input})$ $= 0.93(79/120) + 0.87(41/120)$ $= 0.93 * 0.66 + 0.87 * 0.34$ $= 0.91$

Results:

Accuracy: 0.91

	Precision	Recall	F1-score	Support
0	0.96	0.90	0.93	79
1	0.83	0.93	0.87	41
macro avg	0.89	0.91	0.90	120
weighted avg	0.91	0.91	0.91	120