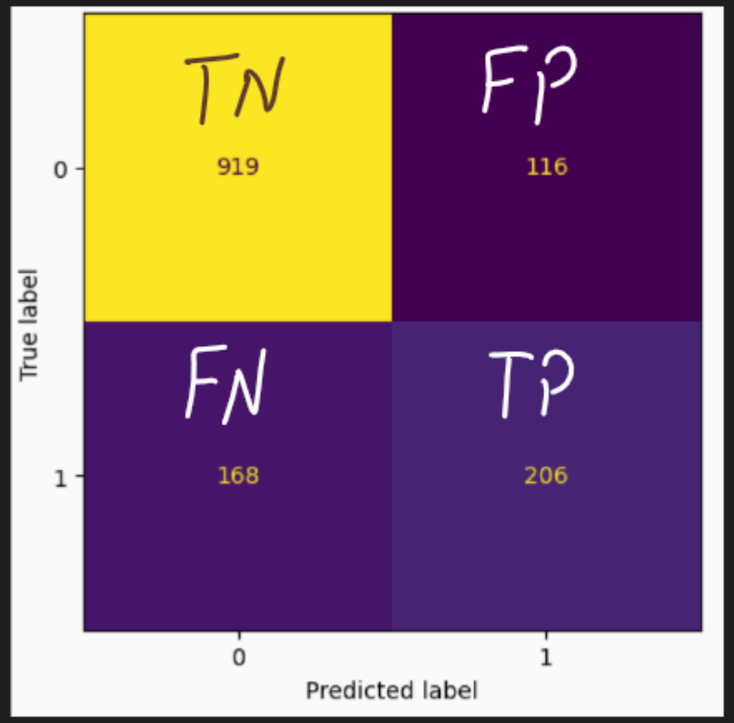


Confusion Matrix

Classification Metrics

① $\text{Accuracy} = \frac{TN + TP}{TN + TP + FN + FP}$ $\quad \frac{\sum(y == \hat{y})}{len(y)}$

② ↑ $\text{Recall} = \frac{TP}{TP + FN\downarrow}$ $\quad \frac{TP}{\text{All Actual +ve class}}$

③ ↑ $\text{precision} = \frac{TP}{TP + FP\downarrow}$ $\quad \frac{TP}{\text{All pred +ve class}}$

Cost(FN) Cost(FP)

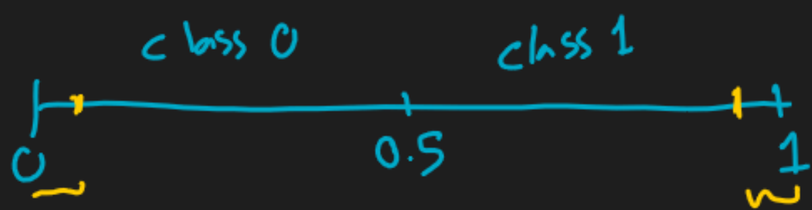
Recall — Tradoff — Precision

* Focus on FN: Examples: Churn - Covid-19

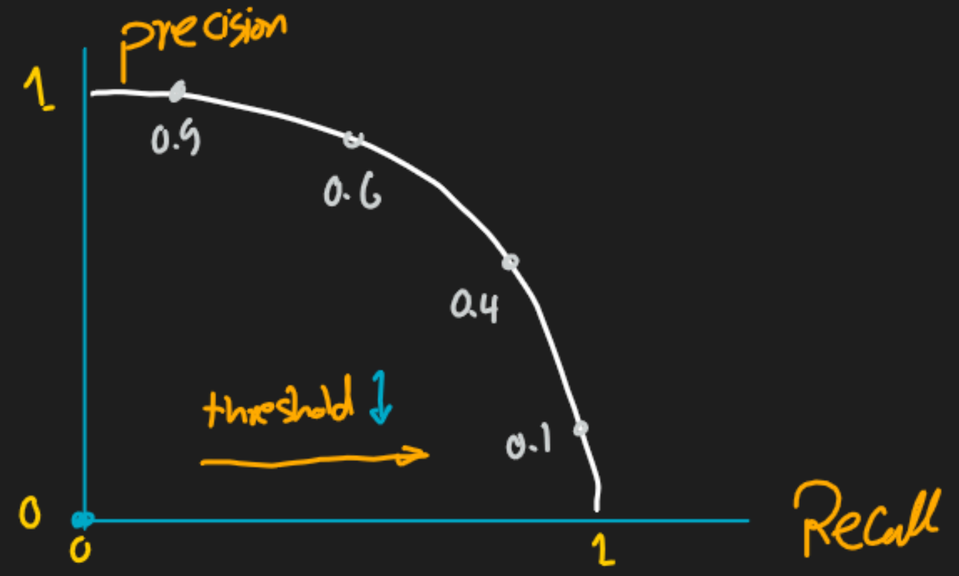
  - threshold << 0.5
  - FN ↓ → Recall ↑

* Focus on Fp: Example: Safe videos for kids

  - Threshold >> 0.5
  - Fp ↓ → precision ↑



FN≈0 → Recall ≈1



Fp≈0 → precision ≈1

precision-Recall Curve

* Cost (FN) ≈ Cost (FP) → precision, Recall, Balance

→ Harmonic Mean of precision & Recall = $\frac{2}{\frac{1}{\text{precision}} + \frac{1}{\text{Recall}}}$

$$\text{F1\_Score} = 2 \times \frac{\text{precision} \times \text{Recall}}{\text{precision} + \text{Recall}}$$

$$F_\beta = (1 + \beta^2) \cdot \frac{\text{Precision} \cdot \text{Recall}}{(\beta^2 \cdot \text{Precision} + \text{Recall})}$$

β=1 → F.2 Score (More weight for recall)
β=0.5 → F.0.5 Score (More weight for precision)

The F-beta score is the weighted harmonic mean of precision and recall, reaching its optimal value at 1 and its worst value at 0.

The `beta` parameter represents the ratio of recall importance to precision importance. `beta > 1` gives more weight to recall, while `beta < 1` favors precision. For example, `beta = 2` makes recall twice as important as precision, while `beta = 0.5` does the opposite. Asymptotically, `beta -> +inf` considers only recall, and `beta -> 0` only precision.

Classification Report

	precision	recall	f1-score	support
0	0.85	0.90	0.87	4139
1	0.67	0.56	0.61	1495
accuracy			0.81	5634
macro avg	0.76	0.73	0.74	5634
weighted avg	0.80	0.81	0.80	5634

False negative: 3722
False positive: 417
True negative: 662
True positive: 833

Actual \ pred	0	1
0	TN	FP
1	FN	TP

Class 1

① $prec\_1 = \frac{TP}{TP+fp} = 0.67$

② $Recall\_1 = \frac{TP}{TP+FN} = 0.56$

③ $F1\_score\_1 = Harmonic\ Mean\ (①, ②)$

⑦ $Accuracy = \frac{TP+TN}{TP+TN+fp+FN} = 0.81$

⑧ $\frac{prec\_1 + prec\_0}{2} = 0.76$

⑨ $\frac{prec\_1 \times 1495 + prec\_0 \times 4139}{1495 + 4139}$

Class 0

④ $prec\_0 = \frac{TN}{TN+FN} = 0.85$

⑤ $Recall\_0 = \frac{TN}{TN+fp} = 0.9$

⑥ $F1\_score\_0 = Harmonic\ Mean\ (④, ⑤)$

Multi-class

Micro Avg . . . . .

Actual		0	1	2
	0	TN	FP	TN
	1	FN	TP	FN
	2	TN	FP	TN

pred

Multiclass

↓

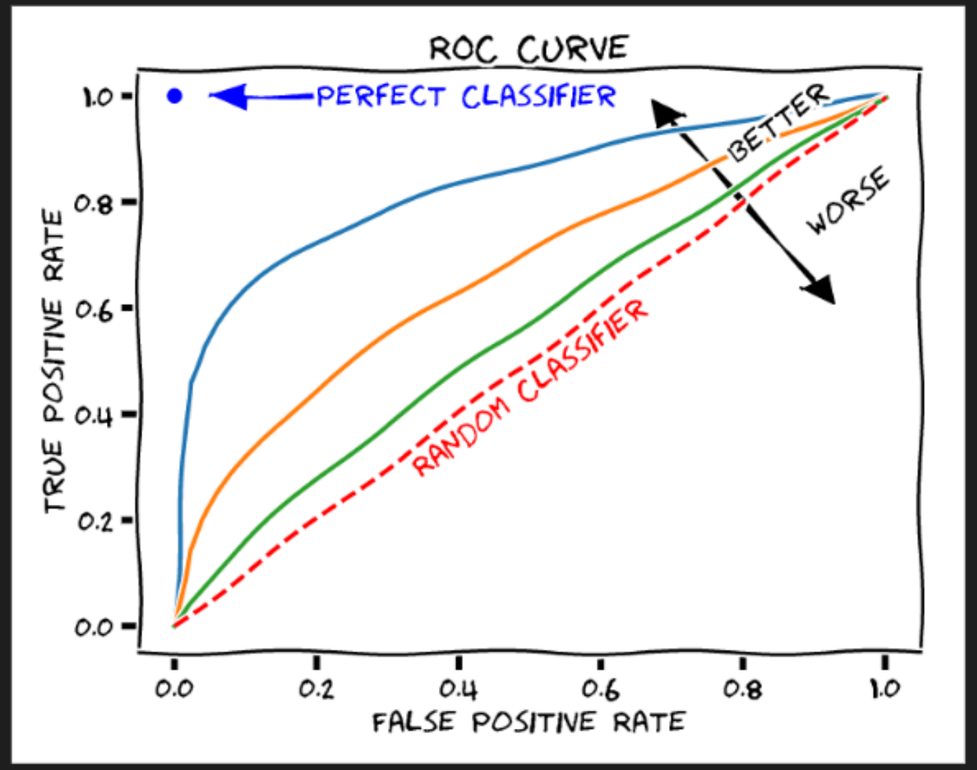
Micro Avg.

Avg. of precision

$$\text{Macro\_Avg} = \frac{\text{prec\_1} + \text{prec\_2} + \text{prec\_0}}{3}$$

$$\text{weighted\_Avg} = \frac{\text{prec\_1} \times n_1 + \text{prec\_2} \times n_2 + \text{prec\_0} \times n_0}{n_1 + n_2 + n_0}$$

$$\text{Micro\_Avg} = \frac{\Sigma TP}{\Sigma TP + \Sigma FP}$$



AUC-ROC-SCore

Actual		0	1
	0	TN	FP
	1	FN	TP
		0	1
		pred	

Specificity

$$FPR = \frac{FP}{FP + TN} = 1 - TNR \quad \text{where } TNR = \frac{TN}{FP+TN}$$

$$TPR = \frac{TP}{TP+FN}$$ Recall (Sensitivity)