

The confusion matrix, decoded.

D8A.ACADEMY/RESOURCES/CONFUSION-MATRIX

THE 2X2

	PREDICTED POSITIVE	PREDICTED NEGATIVE
actually positive	TP · caught it	FN · Type II miss
actually negative	FP · Type I alarm	TN · correctly ignored

HOW TO READ IT

rows	the truth (actual class)
columns	the model call (predicted)
diagonal	where you want the mass

THE FOUR FORMULAS

accuracy	$(TP + TN) / \text{everything}$
precision	$TP / (TP + FP)$
recall	$TP / (TP + FN)$
F1	$2PR / (P + R)$

TYPE I VS TYPE II

Type I · FP	false alarm: convicted an innocent
Type II · FN	the miss: let the guilty walk

WATCH OUT

imbalance	a fat TN cell inflates accuracy
multiclass	NxN grid, same reading rules

→ rows are the truth, columns are your model, mass belongs on the diagonal

WORTH MEMORIZING

01 The 2x2, decoded

Rows are the truth. Columns are the model. The diagonal is success.

	PREDICTED +	PREDICTED -
actually +	TP	FN · Type II
actually -	FP · Type I	TN

→ Every evaluation metric you have ever met is arithmetic on these four cells.

02 Four formulas, one grid

Read them straight off the cells.

accuracy	$(TP + TN) / \text{all}$. Fine only when classes balance.
precision	$TP / (TP + FP)$. Of my alarms, how many real?
recall	$TP / (TP + FN)$. Of the real, how many caught?
F1	harmonic mean. One number, imbalance-aware.

03 Trap: the fat TN cell

A huge true-negative cell makes accuracy look great while the model misses everything rare.

symptom	accuracy 99%, business value zero
the check	read recall on the rare class first
go deeper	sheet 006: precision vs recall vs F1

→ Rare-event problems (fraud, churn, defects) live or die on the FN cell.