

Confusion Matrix

É uma matriz 2x2, onde temos a contagem de valores preditos do modelo. Basicamente essa contagem é dividida em quatro partes, VP, VN, FP, FN, ou seja como abaixo:

		Valores Reais	
		Positivo	Negativo
Previsões	Positivo	Verdadeiro Positivo	Falso Positivo
	Negativo	Falso Negativo	Verdadeiro Negativo

Ela é bem útil por exemplo em problemas relacionados a diagnóstico, pois não queremos que nosso modelo possua muitos falsos negativos, uma vez que o pior caso é prever que o paciente não tem a doença, sendo que no entanto possui. Cada tipo de problema envolvendo classificação pode levar em conta uma parte da matriz como mais importante.

Exemplo de código da matriz, na predição de câncer de mama:

⌵ ▶ ⏪ ⏩ ⌵ ... 🗑

```
1 from sklearn.metrics import confusion_matrix, classification_report
2
3 print(confusion_matrix(y_test, y_pred))
4 print(classification_report(y_test, y_pred))
```

[20] ✓ 0.0s

```
... [[ 62  1]
      [ 3 105]]
```

	precision	recall	f1-score	support
0	0.95	0.98	0.97	63
1	0.99	0.97	0.98	108
accuracy			0.98	171
macro avg	0.97	0.98	0.98	171
weighted avg	0.98	0.98	0.98	171