

# 模型评估

## 1.1 基本概念

- 真阳性 (True Positive, TP) : 预测为正, 实际也为正
- 假阳性 (False Positive, FP) : 预测为正, 实际为负
- 假阴性 (False Negative, FN) : 预测为负、实际为正
- 真阴性 (True Negative, TN) : 预测为负、实际也为负
- 假阳率 (False Positive Rate, FPR) : 原本是错的预测为对的比例  
(越小越好, 0为理想状态)

$$FPR = \frac{FP}{FP+TN} \quad FPR = \frac{FP}{FP+TN}$$

- 真阳率 (True Positive Rate, TPR) : 原本是对的预测为对的比例  
(越大越好, 1为理想状态)

$$TPR = \frac{TP}{TP+FN} \quad TPR = \frac{TP}{TP+FN}$$

- 精确率 (查准率, Precision) : 预测为对的当中, 原本为对的比例  
(越大越好, 1为理想状态)

$$Precision = \frac{TP}{TP+FP} \quad Precision = \frac{TP}{TP+FP}$$

- 召回率 (查全率, Recall) : 原本为对的当中, 预测为对的比例 (越大越好, 1为理想状态)

$$Recall = \frac{TP}{TP+FN} \quad Recall = \frac{TP}{TP+FN}$$

- 准确率 (Accuracy) : 预测对的 (包括原本是对预测为对, 原本是错的预测为错两种情形) 占整个的比例 (越大越好, 1为理想状态)

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

## 1.2 F-score

F分数是对准确率和召回率做一个权衡, 公式为:

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

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

$\beta$  用于调和Precision和Recall的重要性, 当  $\beta$  为1时同等重要, 称为F1-score。

- 微平均Micro-F1: 计算出所有类别总的Precision和Recall, 然后计算F1

- 宏平均Macro-F1：计算出每一个类的Precision和Recall后计算F1，最后将F1平均

```
from sklearn.metrics import f1_score
```

```
ma_f1 = f1_score(labels, y_pred, average='macro')
```

```
mi_f1 = f1_score(labels, y_pred, average='micro')
```

```
print(ma_f1, mi_f1)
```

```
"""
```

```
0.786153332589687,0.7858
```

```
"""
```