

# precision & recall

---

在分类问题中，特别是在二元分类中，precision（精确率）和recall（召回率）是衡量模型性能的两个重要指标。

主要是二分类模型中使用

首先模型将样本分为Positive和Negative两类，表示正类和负类，代表着有这两种标签分类结果。

分类完之后，我们得到了分类结果，有True和False两种结果，代表分类正确和错误分类，具体情况如下图所示

- **TP (True Positives)**: 真正例，模型正确地将正例预测为正例。
- **TN (True Negatives)**: 真负例，模型正确地将负例预测为负例。
- **FP (False Positives)**: 假正例，模型错误地将负例预测为正例。有时也被称作 Type I 错误。
- **FN (False Negatives)**: 假负例，模型错误地将正例预测为负例。有时也被称作 Type II 错误。

## Precision（精确率）

精确率是指在所有被模型预测为正类的样本中，实际上为正类的样本所占的比例。它反映了模型对正类预测的准确性。数学上可以表示为：

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

其中，TP（True Positives）是真正类的数量，即模型正确预测为正类的样本数；FP（False Positives）是假正类的数量，即模型错误预测为正类的样本数。

这里的精确率，分子为True Positive，分母直接为Positive的总和

理解下来就是我的模型将一堆输入（包括正例和负例）分类成了正例，现在我就要判断我模型的精确率。

## Recall（召回率）

召回率是指在所有实际为正类的样本中，被模型正确预测为正类的样本所占的比例。它反映了模型捕捉正类样本的能力。数学上可以表示为：

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

其中，FN（False Negatives）是假负类的数量，即模型错误预测为负类的正类样本数。

召回率的分子还是TP则代表正确判断正类样本为正类的量，分母为正确将正类样本判断为正类的量加上错误将正类样本预测成负类的量。

理解下来就是我的模型将一堆正例输入分类成了正例和反例，现在我要判断模型的召回率（查全率）

记忆：

1. 精确率的输入为有正有负，但是最后结果都是判断为正
2. 召回率的输入全部都是正，但是最后结果判断有正有负
3. 这二者的分子都相同，分母就是输入

## 实例

假设我们的任务是从一系列的猫咪图片中识别出那些有蓝色瞳孔的猫咪。

我们有一个自动化的机器学习模型来进行这项任务，并且我们对该模型进行了一次性能评估。在评估过程中，我们手头有100张猫咪的图片，其中有30张是蓝色瞳孔的猫咪。我们的模型对这100张图片进行了分类，结果如下：

- 模型识别出了25张图片为蓝色瞳孔的猫咪。
- 实际上，这25张被识别的图片中有20张确实是蓝色瞳孔的猫咪。
- 而实际上有蓝色瞳孔的30张图片中，有10张被模型错过了，没有被正确识别。

现在我们来计算精确率和召回率：

- **精确率 (Precision)**：模型预测为蓝色瞳孔的猫咪中，实际上确实是蓝色瞳孔的概率是多少？

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}} = \frac{20}{20+5} = \frac{20}{25} = 0.8$$

因此，精确率是80%。

- **召回率 (Recall)**：所有真正有蓝色瞳孔的猫咪中，模型成功识别出来的概率是多少？

$$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}} = \frac{20}{20+10} = \frac{20}{30} \approx 0.67$$

因此，召回率是大约67%。

在这个例子中，虽然模型有一个相当高的精确率（80%），意味着其识别出的蓝色瞳孔猫咪大多数是正确的，但召回率（67%）表明模型错过了相当一部分实际为蓝色瞳孔的猫咪，没有被模型识别出来。在现实世界的应用中，我们可能需要根据任务的重要性来决定是要优先提高精确率还是召回率。

# 关于准确率和精确率的区别

---

## 准确率 (Accuracy)

准确率是所有分类正确的样本（正例和负例） 占有所有样本总数的比例。它是一个衡量模型整体性能的指标。准确率的公式是：

$$\text{Accuracy} = \frac{\text{True Positives (TP)} + \text{True Negatives (TN)}}{\text{Total Number of Samples}}$$

在所有分类任务中，一个高准确率的模型能够正确地分类出大部分的正例和负例。

分子为T，分母为所有样本

当正例和负例相差量十分巨大的时候，往往Accuracy不再那么管用，我们需要的就是精确率precision

## F1分数

---

F1分数 (F1 Score) 是精确率和召回率的调和平均数，它是两者的综合度量，用于评估分类模型的准确性，特别是在数据类别分布不均衡的情况下。F1分数的取值范围为0到1，值越高表示模型性能越好。

F1分数的计算公式为：

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

如果我们回到之前的猫咪瞳孔颜色分类例子，我们可以计算F1分数如下：

已知精确率 (Precision) 是80% (或0.8) ， 召回率 (Recall) 是大约67% (或0.67) ，  
F1分数将是：

$$F1 = 2 \times \frac{0.8 \times 0.67}{0.8 + 0.67} \approx 0.73$$

### F1分数的范围是从0到1

F1分数的范围是0到1， 其中：

- **F1分数为1**：表示模型的精确率和召回率都是完美的。
- **F1分数为0**：表示模型的精确率和召回率至少有一个是非常差的。

F1分数和精确率与召回率之间的关系如下：

- 如果精确率和召回率都高，F1分数也会高。
- 如果精确率和召回率之一很低，即使另一个指标很高，F1分数也会受到影响，因为调和平均数会向较小的数值倾斜。
- 在精确率和召回率相等时，F1分数等于它们的值。
- 如果我们提高精确率而牺牲召回率，或者提高召回率而牺牲精确率，F1分数可能会下降，因为它寻求两者之间的平衡。