

如何解释召回率与准确率? - 知乎

笔记本:	机器学习		
创建时间:	2017/12/13 10:44	更新时间:	2018/6/24 8:36
作者:	beyourselfwb		
URL:	https://www.zhihu.com/question/19645541		

算法

如何解释召回率与准确率?

关注者

875

被浏览

131904

32 个回答

怕什么真理无穷，进一寸有一寸的欢喜

总结之后，我的理解如下：

accuracy --- 准确率

说白了就是：你的预测有多少是对的？公式如下：

预测与样本相符的

accuracy = $\frac{\text{预测与样本相符的}}{\text{all}}$

precision --- 精确率 (针对预测结果)

预测为正的样本中，有多少是真正的正样本？

TP

precision = $\frac{\text{TP}}{\text{所有预测的正例}}$

recall --- 召回率 (针对原样本)

正例里你的预测覆盖了多少？

TP

recall = $\frac{\text{TP}}{\text{样本所有正例}}$

$$P = \frac{\text{正确识别的实体数}}{\text{识别的实体数}} \quad (17)$$

$$R = \frac{\text{正确识别的实体数}}{\text{样本的实体数}} \quad (18)$$

$$F_1 = \frac{2 \times \text{准确率} \times \text{召回率}}{\text{准确率} + \text{召回率}} \quad (19)$$

其实就是分母不同，一个分母是预测为正的样本数，另一个是原来样本中所有的正样本数。

在信息检索领域，精确率和召回率又被称为**查准率**和**查全率**，

查准率 = 检索出的相关信息量 / 检索出的信息总量

查全率 = 检索出的相关信息量 / 系统中的相关信息总量

二.举个栗子

准确率(accuracy) = 预测对的/所有 = $(TP+TN)/(TP+FN+FP+TN) = 70\%$

精确率(precision) = $TP/(TP+FP) = 80\%$

召回率(recall) = $TP/(TP+FN) = 2/3$

纠正下，**精确率**(precision)和**准确率**(accuracy)是不一样的，题主问的应该是精确率与召回率之间的差别。

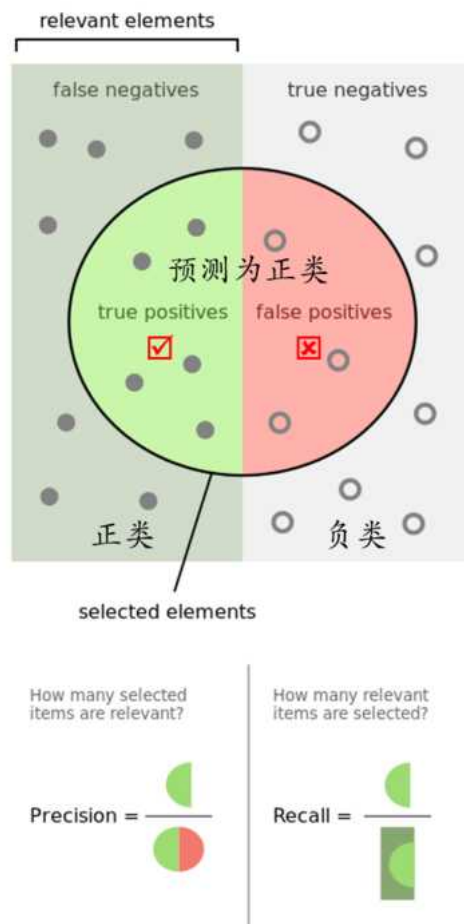
刚开始接触这两个概念的时候总搞混，时间一长就记不清了。

实际上非常简单，**精确率**是针对我们**预测结果**而言的，它表示的是预测为正的样本中有多少是真正的正样本。那么预测为正就有两种可能了，一种就是把正类预测为正类(TP)，另一种就是把负类预测为正类(FP)，也就是

$$P = \frac{TP}{TP + FP}$$

而**召回率**是针对我们原来的**样本**而言的，它表示的是样本中的正例有多少被预测正确了。那也有两种可能，一种是把原来的正类预测成正类(TP)，另一种就是把原来的正类预测为负类(FN)。

$$R = \frac{TP}{TP + FN}$$



假设我们手上有60个正样本，40个负样本，我们要找出所有的正样本，系统查找出50个，其中只有40个是真正的正样本，计算上述各指标。

- TP: 将正类预测为正类数 40
- FN: 将正类预测为负类数 20
- FP: 将负类预测为正类数 10
- TN: 将负类预测为负类数 30

除此之外，还有ROC曲线，PR曲线，AUC等评价指标，可以参见我的博客：[机器学习性能评估指标](#)
编辑于 2016-11-01



知乎用户

博客: [muyefeifei.com](#)

这篇文章讲的挺简单的，很容易理解，链接：[推荐系统评测指标](#)

准确率和**召回率**是广泛用于信息检索和统计学分类领域的两个度量值，用来评价结果的质量。其中精度是检索出相关文档数与检索出的文档总数的比率，衡量的是检索系统的**查准率**；召回率是指检索出的相关文档数和文档库中所有的相关文档数的比率，衡量的是检索系统的**查全率**。

一般来说，Precision就是检索出来的条目（比如：文档、网页等）有多少是准确的，Recall就是所有准确的条目有多少被检索出来了。

正确率、召回率和 F 值是在鱼龙混杂的环境中，选出目标的重要评价指标。不妨看看这些指标的定义先：

1. 正确率 = 提取出的正确信息条数 / 提取出的信息条数

2. 召回率 = 提取出的正确信息条数 / 样本中的信息条数

两者取值在0和1之间，数值越接近1，查准率或查全率就越高。

3. $F\text{值} = \text{正确率} * \text{召回率} * 2 / (\text{正确率} + \text{召回率})$ (F 值即为正确率和召回率的调和平均值)

不妨举[这样一个例子](#)：某池塘有1400条鲤鱼，300只虾，300只鳖。现在以捕鲤鱼为目的。撒一大网，逮着了700条鲤鱼，200只虾，100只鳖。那么，这些指标分别如下：

正确率 = $700 / (700 + 200 + 100) = 70\%$

召回率 = $700 / 1400 = 50\%$

$F\text{值} = 70\% * 50\% * 2 / (70\% + 50\%) = 58.3\%$

不妨看看如果把池子里的所有的鲤鱼、虾和鳖都一网打尽，这些指标又有何变化：

正确率 = $1400 / (1400 + 300 + 300) = 70\%$

召回率 = $1400 / 1400 = 100\%$

$F\text{值} = 70\% * 100\% * 2 / (70\% + 100\%) = 82.35\%$

由此可见，正确率是评估捕获的成果中目标成果所占得比例；召回率，顾名思义，就是从关注领域中，召回目标类别的比例；而F值，则是综合这二者指标的评估指标，用于综合反映整体的指标。

当然希望检索结果Precision越高越好，同时Recall也越高越好，但事实上这两者在某些情况下有矛盾的。比如极端情况下，我们只搜索出了一个结果，且是准确的，那么Precision就是100%，但是Recall就很低；而如果我们把所有结果都返回，那么比如Recall是100%，但是Precision就会很低。因此在不同的场合中需要自己判断希望Precision比较高或是Recall比较高。如果是做实验研究，可以绘制**Precision-Recall曲线**来帮助分析。

发布于 2015-02-14

