# 一 评价标准有哪些

## 1.1 分类问题性能指标---准确率,精确率,召回率

准确率: 对于测试集,分类器能正确分类样本数与总样本之比.

TP: true positive, 正类判为正类.

FN: false negative, 正类判为负类.

FP: flase positive, 负类判为正类.

TN: true negtive, 负类判为负类.

>> 记忆法:

True和false表示是否正确判别.

Positive和negetive是被判别成什么类.

TP,TN是判定正确的,因此是正判正,负判负.

FP,FN是判定错误的, 因此是负判正,正判负.

精确率: precision = TP/(TP+FP)

正例结果中,实际为正例的比率.

召回率: recall = TP/(TP+FN)

正确被分类的结果(TP)占应该被分类的比率.

## 1.2 实例

### 1.2.1 例一

有10000张图片,其中一张色情图片, 色情图片为正样本, 分类器分类出2张色情图片, 其中一张是真正的色情图片.

问该分类器的准确率,精确率,召回率.

TP = 1

FP = 1

FN = 0

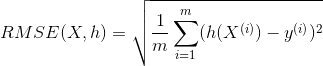
准确率: (1+9998)/10000 = 99.99%

精确率: TP/(TP+FP) = 1/2=50%

召回率: TP/(TP+FN) = 1 = 100%

## 1.3 回归问题的性能指标

回归问题是典型的监督学习的一类算法，目标是通过有标签的一组数据训练出一个线性模型，对于输入能够产生一个预测值。回归问题的典型指标是均方根误差（RMSE），它测量的是系统预测误差的标准差，例如RMSE等于50000，意味着68%的系统封预测值位于实际值的5000之内，95%的预测值位于实际值的100000之内（一个特征通常符合高斯分布，即满足“68-95-99.7”规则：大约68%的值落在1内，95%的值落在2内，99.7%的值落在3内，这里的σ等于50000），RMSE的计算公式如下：



其中：

1. m是RMSE数据集中实例的个数。
2. 是数据集第i个实例的所有特征值（不包含标签）的向量，y是它的标签（这个实例的输出值）。



1. h是系统的预测函数，也称为假设。当系统收到一个实例的特征向量https://private.codecogs.com/gif.latex?X%5E%7B%28i%29%7D，就会输出这个实例的一个预测值\tilde{y}=h(X)

RMSE(X,h)是使用假设h在样本上测量的损失函数。