

1、 FN：False Negative,被判定为负样本，但事实上是正样本。

2、 FP：False Positive,被判定为正样本，但事实上是负样本。

3、TN：True Negative,被判定为负样本，事实上也是负样本。

4、TP：True Positive,被判定为正样本，事实上也是证样本。

5、precesion：查准率，即在检索后返回的结果中，真正正确的个数占整个结果的比例。

precesion = TP/(TP+FP) 即，检索结果中，都是你认为应该为正的样本（第二个字母都是P），但是其中有你判断正确的和判断错误的（第一个字母有T ，F）。

6、 recall：查全率，即在检索结果中真正正确的个数 占整个数据集（检索到的和未检索到的）中真正正确个数的比例。

recall = TP/(TP+FN)即，检索结果中，你判断为正的样本也确实为正的，以及那些没在检索结果中被你判断为负但是事实上是正的（FN）。

7、F-Measure是Precision和Recall加权调和平均

P和R指标有时候会出现的矛盾的情况，这样就需要综合考虑他们，最常见的方法就是F-Measure（又称为F-Score）。

8、MCC 马修斯相关系数

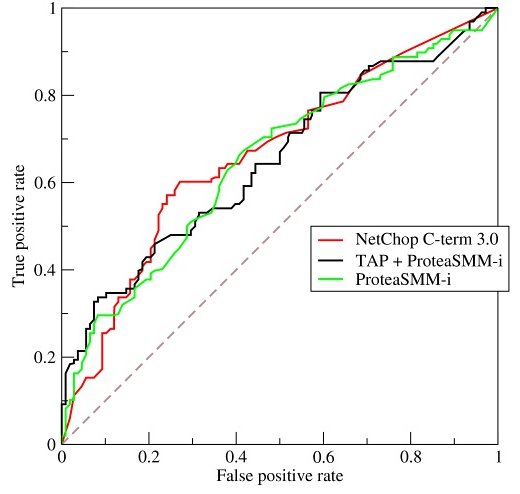
衡量不平衡数据集的指标比较好，公式是：

C:\Users\Administrator.PCOS-1408051144\AppData\Roaming\Tencent\Users\1031545274\QQ\WinTemp\RichOle\1[U~U%WS1{}FM)9V%_9FOHO.png

9、ROC和AUC

ROC曲线的横坐标为false positive rate（FPR），纵坐标为 true positive rate（TPR）

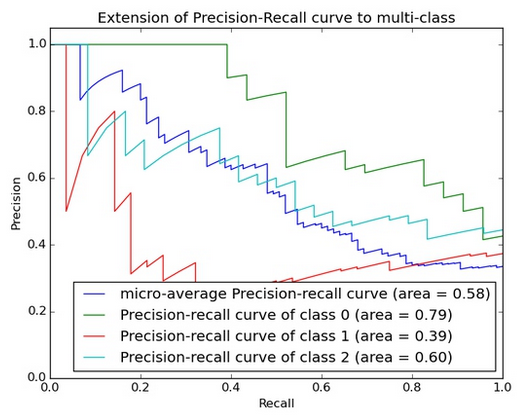
当测试集中的正负样本的分布变化的时候，ROC曲线能够保持不变。根据每个测试样本属于正样本的概率值从大到小排序，依次将 “Score”值作为阈值threshold，当测试样本属于正样本的概率 大于或等于这个threshold时，认为它为正样本，否则为负样本。



计算出ROC曲线下面的面积，就是AUC的值。 介于0.5和1.0之间，越大越好。

10、PRC曲线

在正负样本分布得极不均匀(highly skewed datasets)的情况下，PRC比ROC能更有效地反应分类器的好坏。



11、kappa statics

Kappa值即内部一致性系数(inter-rater,coefficient of internal consistency)，是作为评价判断的一致性程度的重要指标。取值在0～1之间。Kappa≥0.75两者一致性较好；0.75>Kappa≥0.4两者一致性一般；Kappa<0.4两者一致性较差。

这个参数是把分类器与随机分类器作比较得出的一个对分类器的评价值。那么0.186是怎么计算出来的呢.

12、Mean absolute error 和 Root mean squared error

平均绝对误差和均方根误差，用来衡量分类器预测值和实际结果的差异，越小越好。

13、Relative absolute error 和 Root relative squared error

相对绝对误差和相对均方根误差，有时绝对误差不能体现误差的真实大小，而相对误差通过体现误差占真值的比重来反映误差大小。