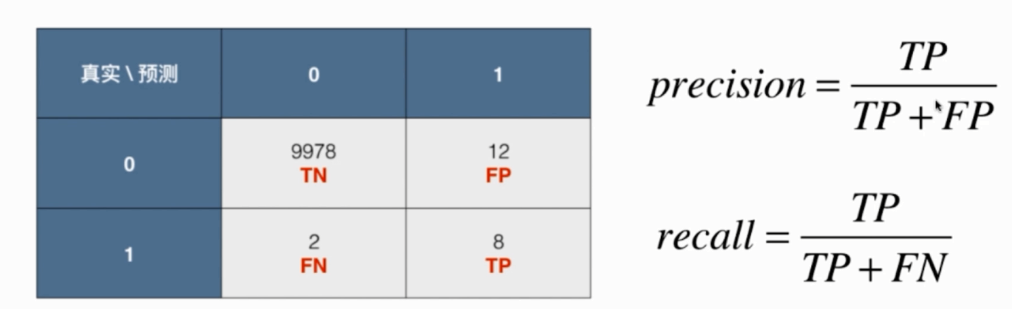
# 精准率和召回率的定义:



举例:

精准率:我们做了100次预测为1的预测,有多少次是准确的.

召回率:实际有100个1的事件发生,我们预测对了多少个.

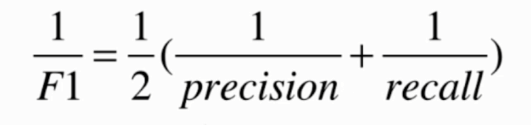
精准率和召回率哪个更重要需要视具体场景而定,

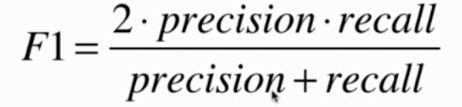
比如如果是绝症的预测,我们希望有病的人一定预测出来,没病的我们预测错了部分没有很大关系,复查可以排除.召回率重要性大于精确率.

如果是股票预测,我们希望预测出来的涨的股票是准确的,当我们把涨的股票预测成跌的股票影响不大,只是错了部分赚钱的机会而已,但是预测出来的结果一定赚钱就OK了,此时精准率就比召回率重要.

当然还有些模型需要同时考虑精准率和召回率,这种模型用f1 score来进行评判.

# F1 score



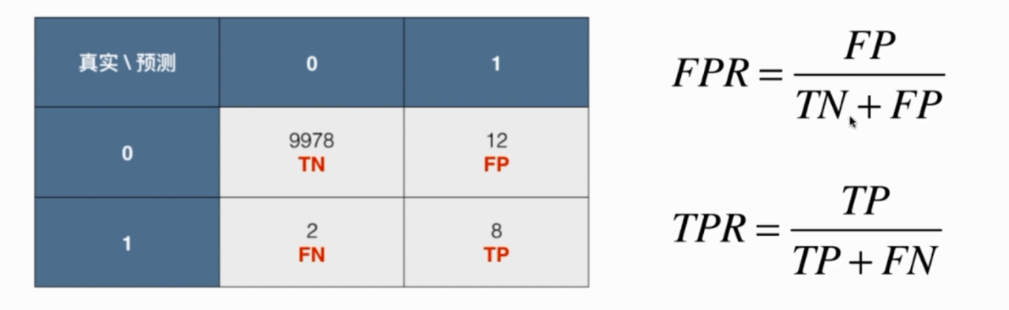


F1 score为精准率和召回率的调和平均值,调和平均值相比算术平均值的特点为,当有一个参数特别小的时候整个调和平均值也会特别小.F1 score 用来衡量精准率和召回率都特别重要的模型.

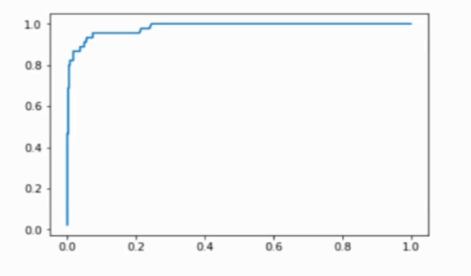
精准率和召回率都重要的模型的好坏的判断,我们可以根据精准率-召回率曲线,ROC曲线来进行筛选.选择面积大的那个.

# ROC 曲线:

FPR(X轴)及TPR(Y轴)在不同判断阈值上的曲线



ROC曲线



在sklearn中求ROC曲线的面积求法:

from sklearn.metrics import roc\_auc\_score

roc\_auc\_score(y\_test,decision\_scores)