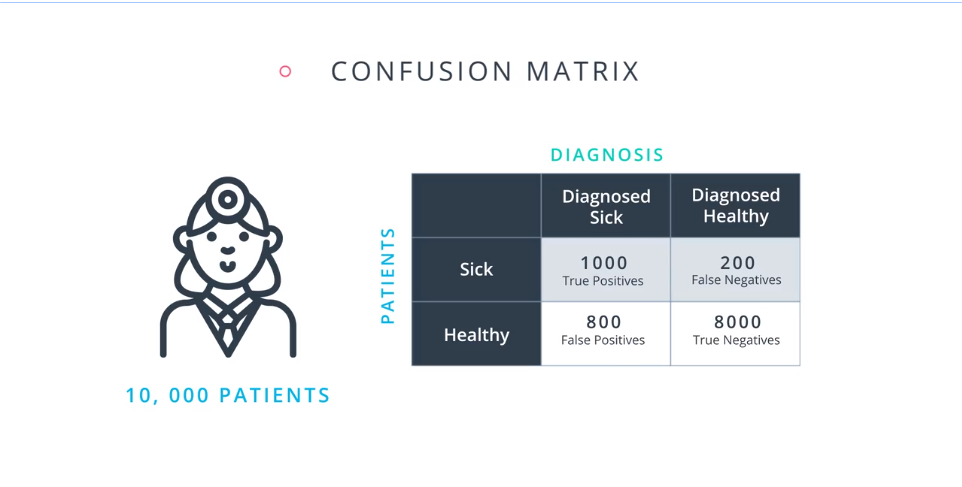
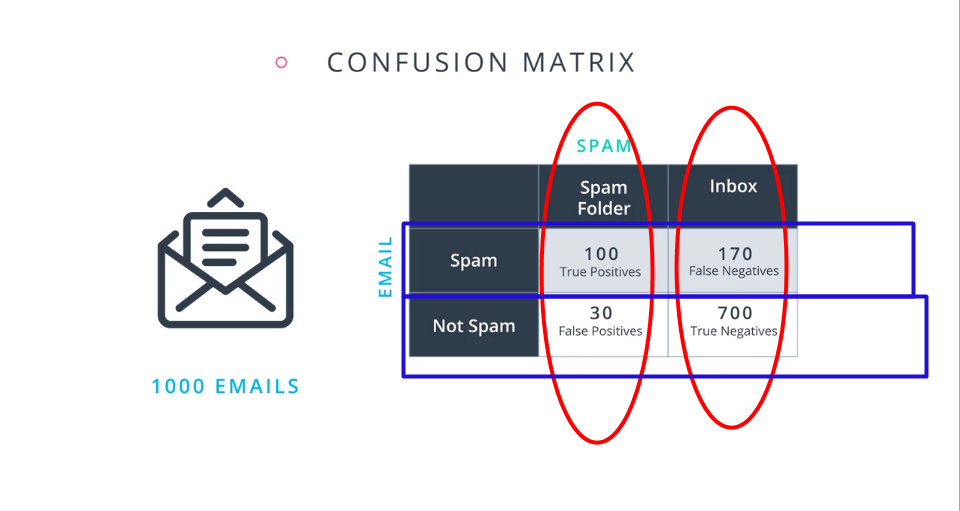
1．混淆矩阵

所谓的混淆矩阵,就是一个用于描述模型性能的表,判断分类好坏程度的方法。只能存储这四个值的表格。

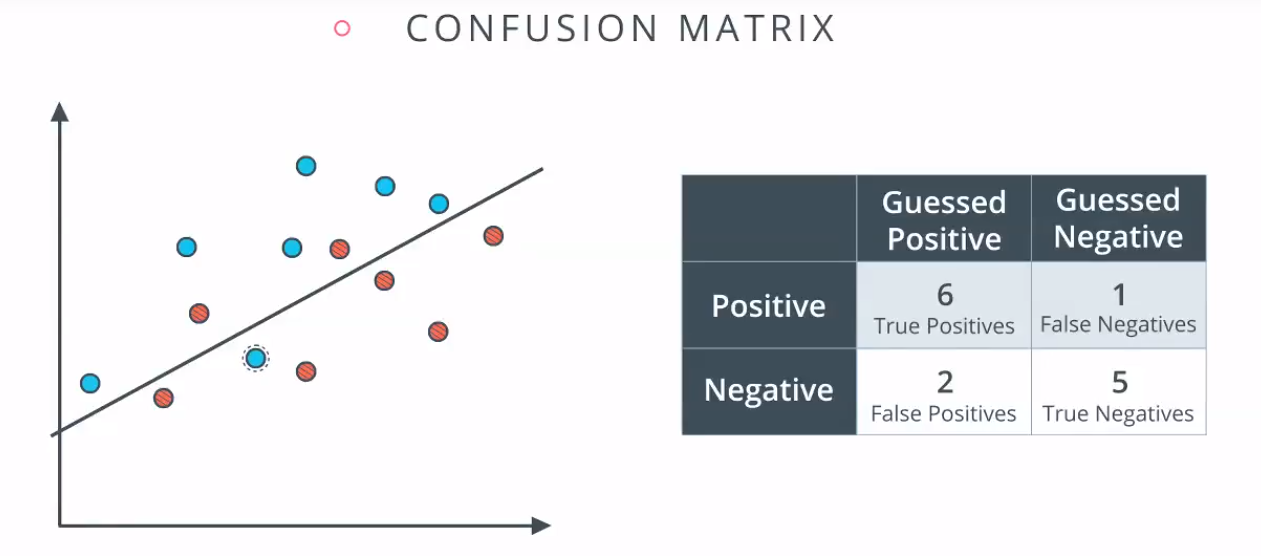
<https://www.jianshu.com/p/0fc8a0b784f1>



Sick为阳，healthy为阴 。

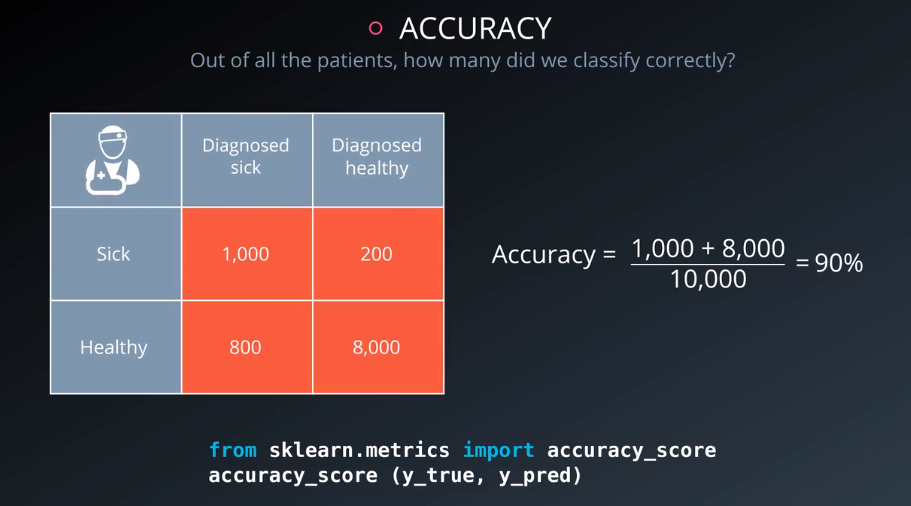


# 2．混淆矩阵Lab

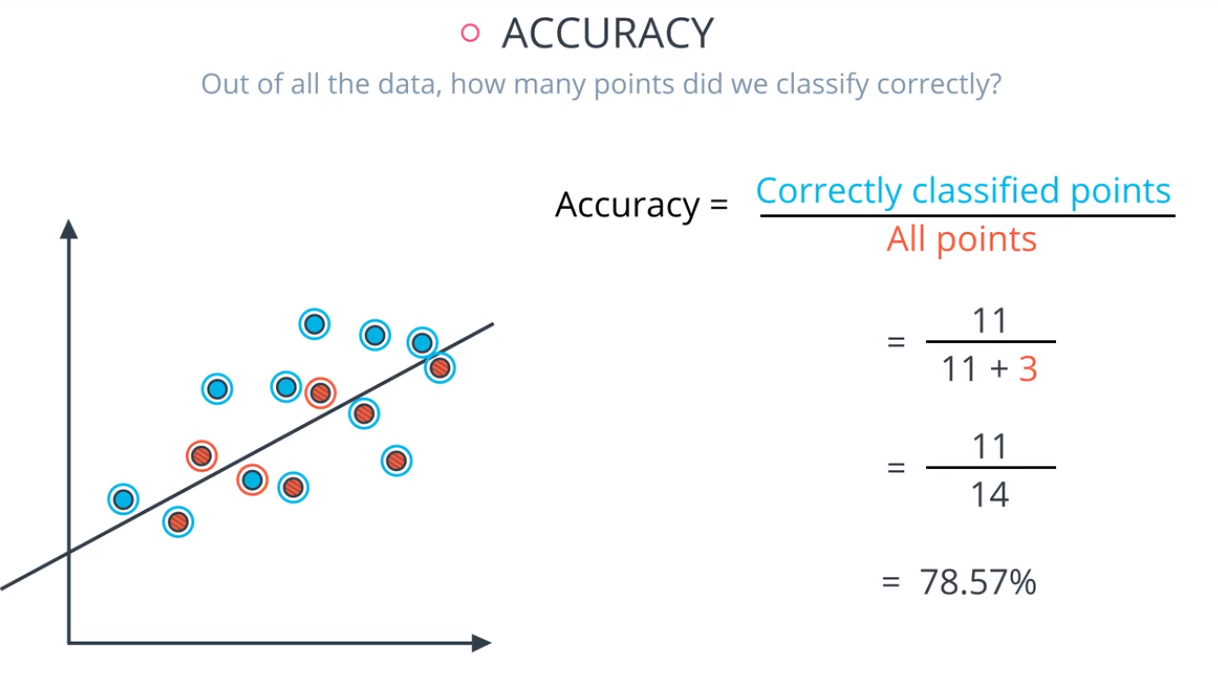


# 3．准确率

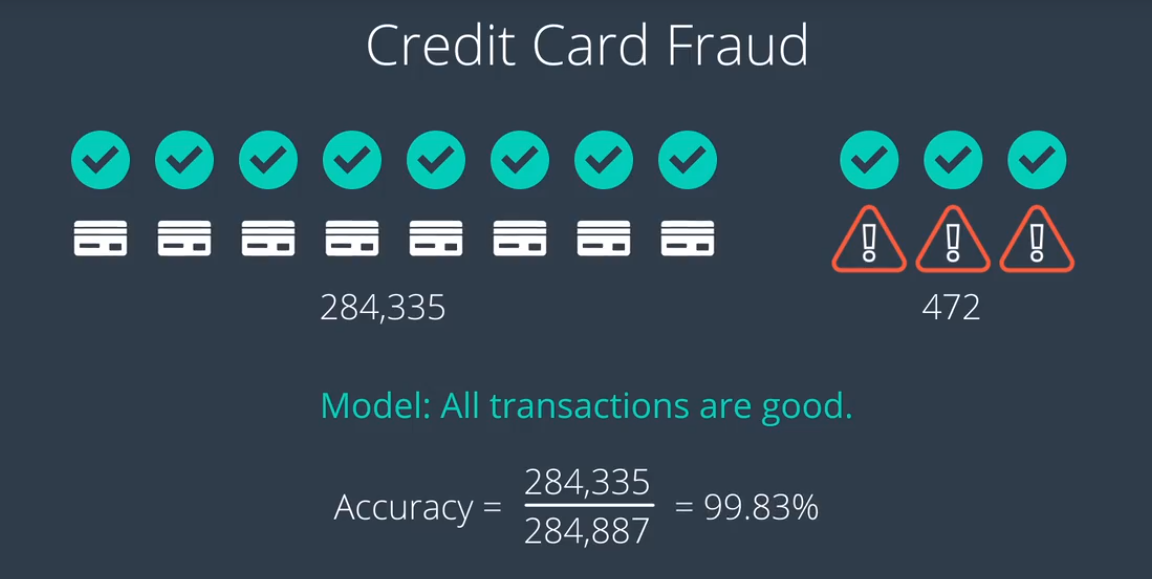
准确率 是 正确分类的点数 与 总点数 的比例。**针对是整体情况。**



# 4．准确率Lab



# 5．准确率不适用的情形



准确率并非始终是最佳指标,例如检测信用卡欺诈行为,其中有些数据是良好的数据.有些是欺诈行为数据 ,如果模型把所有记录都标为良好交易，准确率超过 99% ，但是他并非是好模型，因为该模型没有发现任何不好的数据,而模型的意义本身就是发现欺诈交易。

简单说就是，数据发生偏斜时，准确率不是一个衡量的指标。

# 6. 假阴性与假阳性

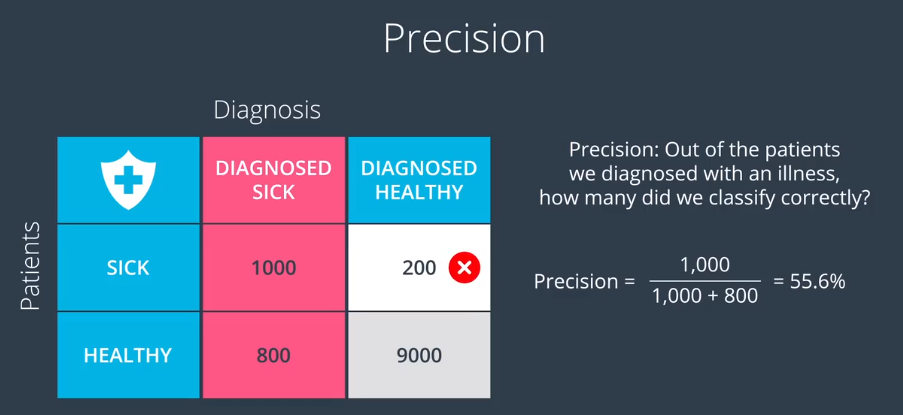
由于医疗模型和电子邮件模型在本质上是完全不同的：

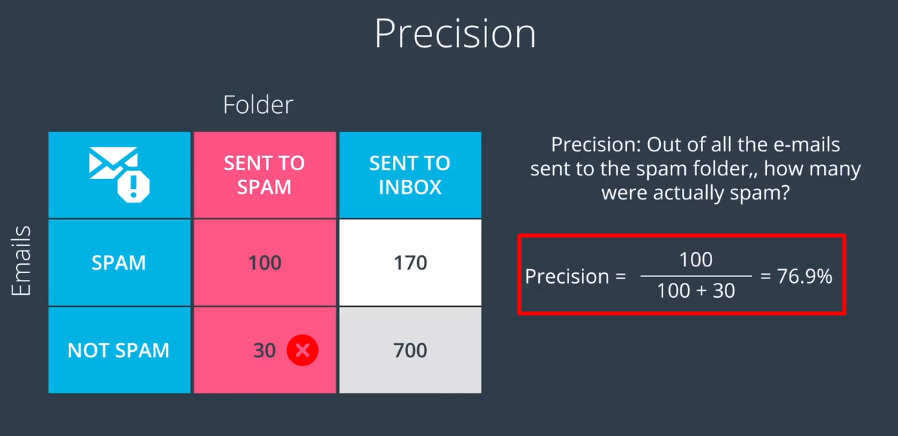
1、医疗模型可以接受假阳性,但不能接受假阴性.(因为我们可以将某人误诊为生病人士)

2、邮件模型可以接受假阴性 但不能接受假阳性.(因为我们不需要找到所有的垃圾邮件)

以下两种指标：精度和召回率，它们将能够衡量我们想衡量的内容，即医疗模型需要是高召回率模型，电子邮件模型需要是高精度模型。

# 8. 精度



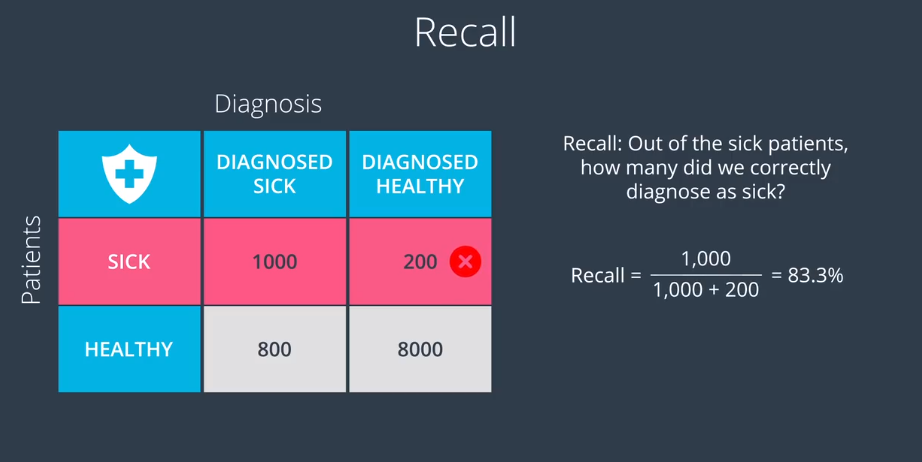


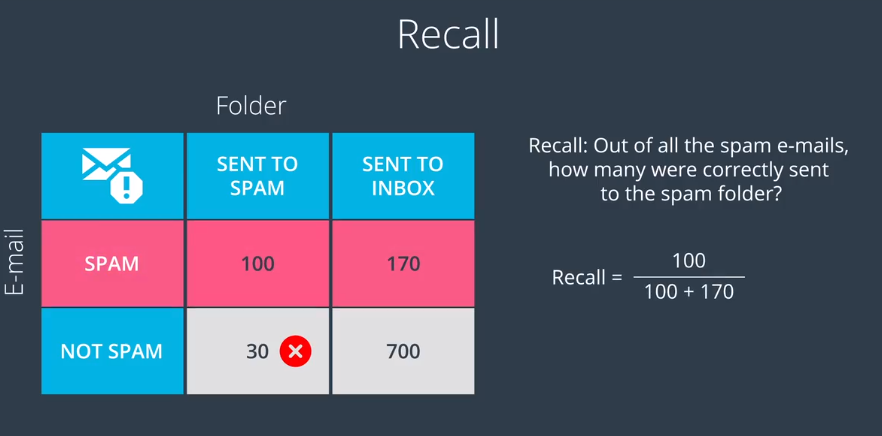
# 9. 召回率

在医疗模型中

精度是诊断为生病的患者当中,有多少位真正生病了？

召回率刚好相反，在生病的患者当中，有多少位正确地诊断为生病了？



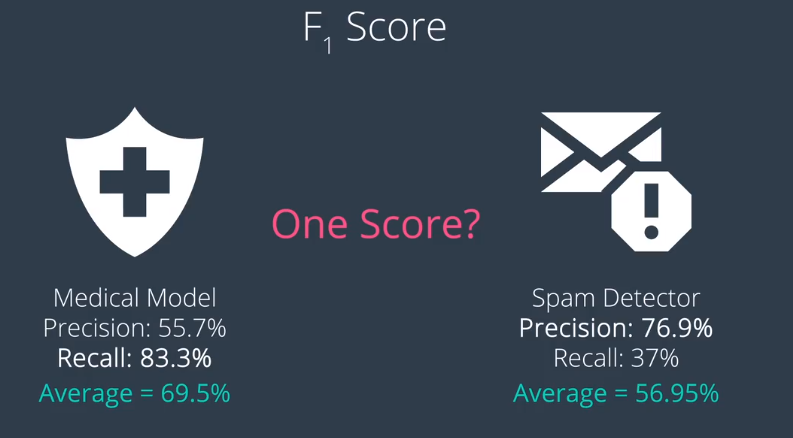


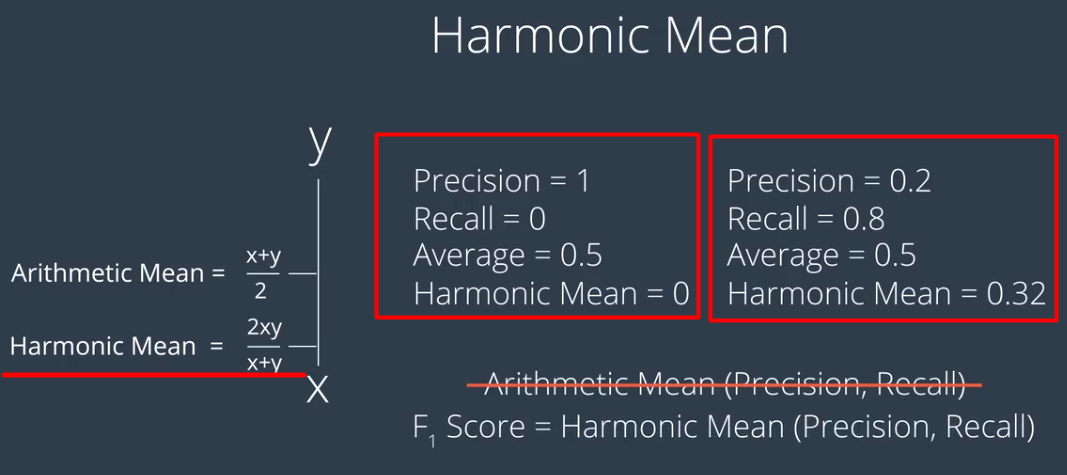
# 10 F1得分,

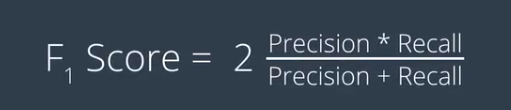
名词解释：Precision:精度、Pecall：召回率，Average：平均数，

Harmonic Mean:调和平均数（F1得分）

F1得分由来，由于精度 和 召回率 这个数据看着不方便，因此需要把它整合为一个数据，为我们提供参考，当数据出现精度很高，召回率很低的时候，平均数是无法看出的，因此才有这种形式进行计算。

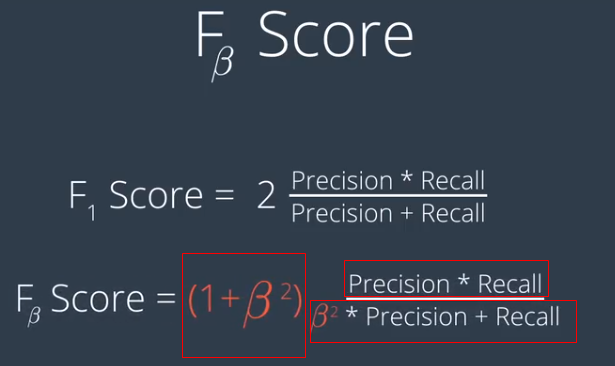




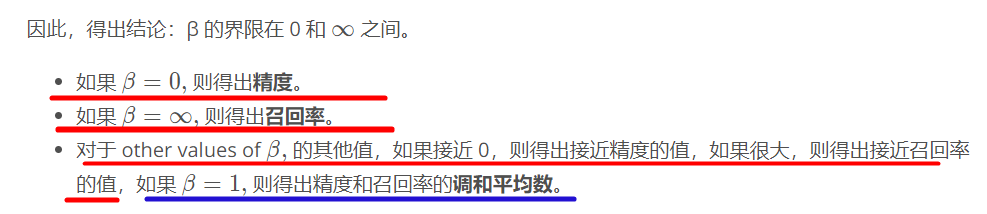




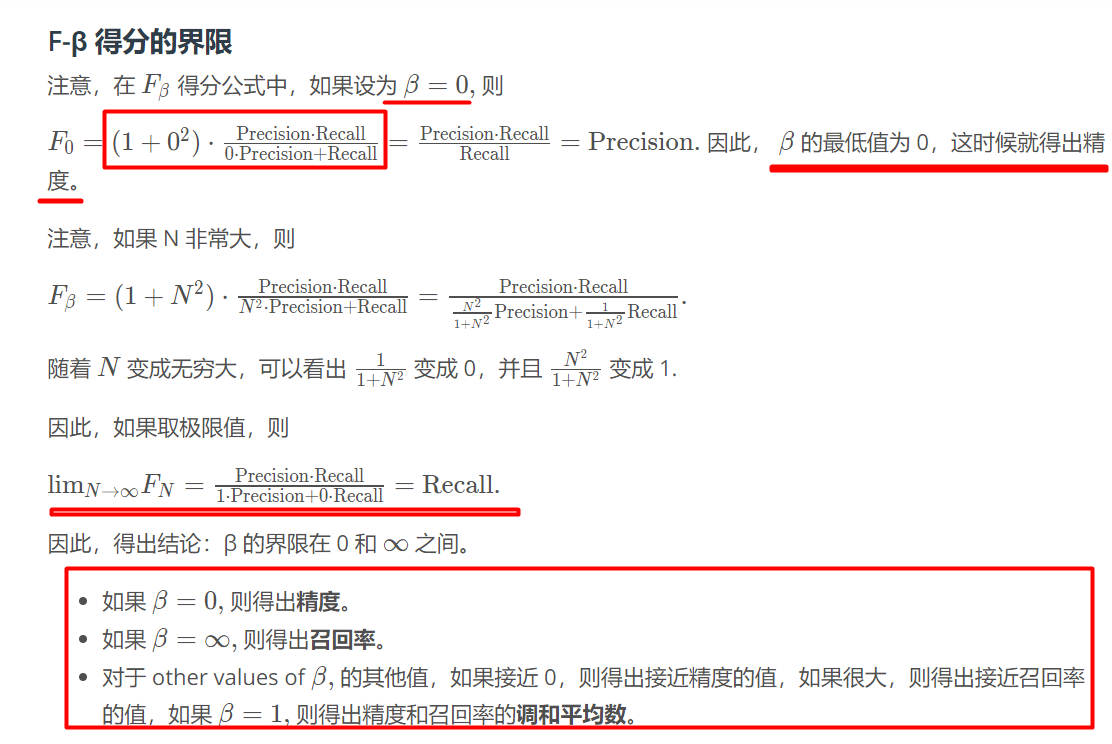
# 11 F-β 得分（F1的分的作用）



## 结论（非常重要）



推到过程



# 总结

我们通过学习一些不同的指标，来了解模型的好坏，例如准确率（真阳+真阴和的比率），精度（假阳性），召回率（假阴性），F1得分（精度与召回率的），F-β得分： β=0 精度， β=∞召回率 β=1 F1得分，其他 在精度与召回率之间。

参考

<https://blog.csdn.net/quiet_girl/article/details/70830796>