

推荐系统常用指标——AUC综述



编码小弟
互联网

关注他

14 人赞同了该文章

首先抛出关于auc的几个问题：

- 1.对数据随机均匀采样，对AUC会有什么影响？
- 2.改变训练集，测试集比例，AUC会有什么变化？
- 3.AUC高，一定代表模型效果好吗，可以举一个例子吗？
- 4.相比PRC与F1，AUC的优势分别是什么？
- 5.AUC适用于什么场景，或为什么推荐中经常会使用AUC作为离线评估指标？
- 6.负样本非常多的情况下，AUC的值会偏高还是偏低

什么是AUC？

AUC的定义是ROC曲线下的面积，实际意义为模型打分时将正例分数排在反例前面的概率。ROC曲线一般都会处于0.5-1之间，所以auc一般是不会低于0.5的，0.5为随机预测的auc。

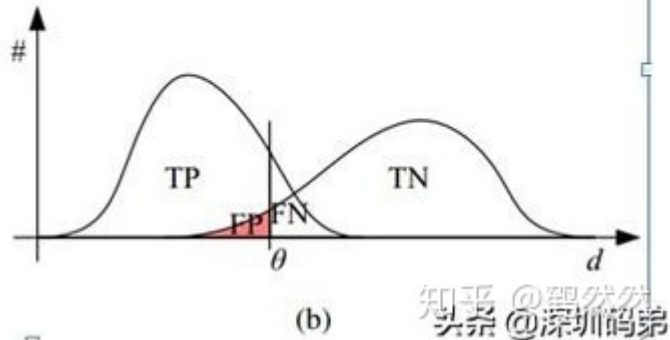
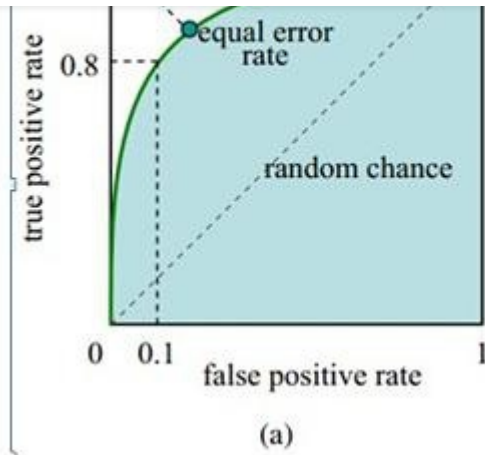
什么是ROC曲线？

ROC中文名为：受试者操作特征（receiver operating characteristic curve），源于二战雷达信号分析技术。

ROC曲线绘制：分别计算模型结果的FPR与TPR，然后将TPR作为纵坐标，TPR作为横坐标作图，便可得到ROC曲线，ROC曲线上的每一个点对应一个阈值。。



知乎

首发于
推荐系统实战

TPR, FPR是什么?

这一块知识非常绕，很容易混淆。

TPR——真正例率： $TP / (TP + FN)$ ，指的是模型预测的正确正例占所有正例的比例，等同于召回率，可以理解为正例的灵敏度。TPR越大，则预测的正例中正例占比越高。

FPR——假负例率： $FN / (TN + FP)$ ，指的是模型预测的错误反例占所有预测反例的比例，等同于反例预测的错误率，也可以理解为模型对负例的特异度。FPR也可以用公式 $(1 - TNR)$ 来表示。FPR越大，则预测的正例中反例越多。

总结：需要FPR越小，TPR越高，则模型越好，所以通常可用ROC曲线最靠近左上角的点作为二分类模型阈值的选择标准。（考虑人群浓度的话则是另一种阈值选法）

如何画ROC曲线?

将设我们预测出 $(z, 0.08), (a, 0.1), (b, 0.2), (c, 0.4), (d, 0.5)$ ，然后我们在预测结果中选择每一个样本的分值作为阈值，比如第一个数据 $(a, 0.1)$ 则分值大于等于0.1的都为正样本，小于0.1的为负样本，然后便根据这些样本算出一组FPR,与TPR值，得到ROC曲线上的一点，对所有测试用例做一遍操作，便可以绘制得到ROC曲线图。

AUC如何计算

方法一：



方法二：

$$AUC = \frac{\sum_{i \in \text{positiveClass}} rank_i - \frac{M(1+M)}{2}}{M \times N}$$

解释：该方法为方法一的进阶版本，首先rank项就是样本按照score值从小到大升序排序，然后只对正样本的序号相加，然后减去正样本在正样本之前的数，结果便是正样本大于负样本的数，然后再除于总的样本数得到的便是AUC值。

直观理解：

A——rank序号：小于该样本分值的样本数。

B—— $M(1+M)/2$:等差数列之和，即将所有正样本排序之和相加（ $1+...+10+...n$ ）

A-B：即将所有小于某正样本得个数去掉，得到的便是小于该正样本的负样本数

$M \times N$ ：所有正负样本两两排序的数之和

最终得到正样本排在负样本之前的概率值。该方法时间复杂度为 $O(n \log n)$

方法三：

直接通过每一个梯形的面积计算整个ROC曲线的面积，该方法目前用的已经不多。

编辑于 2020-07-18

推荐系统 机器学习 深度学习 (Deep Learning)

▲ 赞同 14 ▼ 6 条评论 ➤ 分享 ♥ 喜欢 ★ 收藏 📄 申请转载 ...

