

机器学习算法之 常用的评价指标

目录

- 评价指标的概念
- 分类指标
- 回归指标
- 排序指标
- 总结

评价指标的概念

- 什么是评价指标

- 灯塔之于大海上的轮船
- 星空之于仰望的人类

- 常见的机器学习评价指标有哪些

- Precision、Recall、ROC、AUC、F1、Logloss、CrossEntropy、MAE、MAPE、RMSE、MAP、NDCG

- 如何选择评价指标

- 根据算法的任务类型
- 根据实际业务的关注点

分类指标

- 精确度(Precision)和召回率(Recall)
- ROC曲线和AUC(Area Under Curve)
- 对数损失(LogLoss)

分类指标之 精确率和召回率

混淆矩阵

真实结果	预测结果		
		正 (P)	负 (N)
	正 (P)	TP	FN
	负 (N)	FP	TN

精确率和召回率

— 精确度：选择的正样本中真正样本的概率

$$P(recision) = \frac{TP}{TP + FP}$$

— 召回率：真正样本中预测为正样本的概率

$$R(ecall) = \frac{TP}{TP + FN}$$

准确率和错误率

$$A(ccuracy) = \frac{TP + TN}{TP + TN + FP + FN}$$

F_alpha Score:

$$\frac{1}{F_{\alpha}} = \frac{\alpha^2}{1 + \alpha^2} \frac{1}{R} + \frac{1}{1 + \alpha^2} \frac{1}{P}$$

分类指标之ROC、AUC

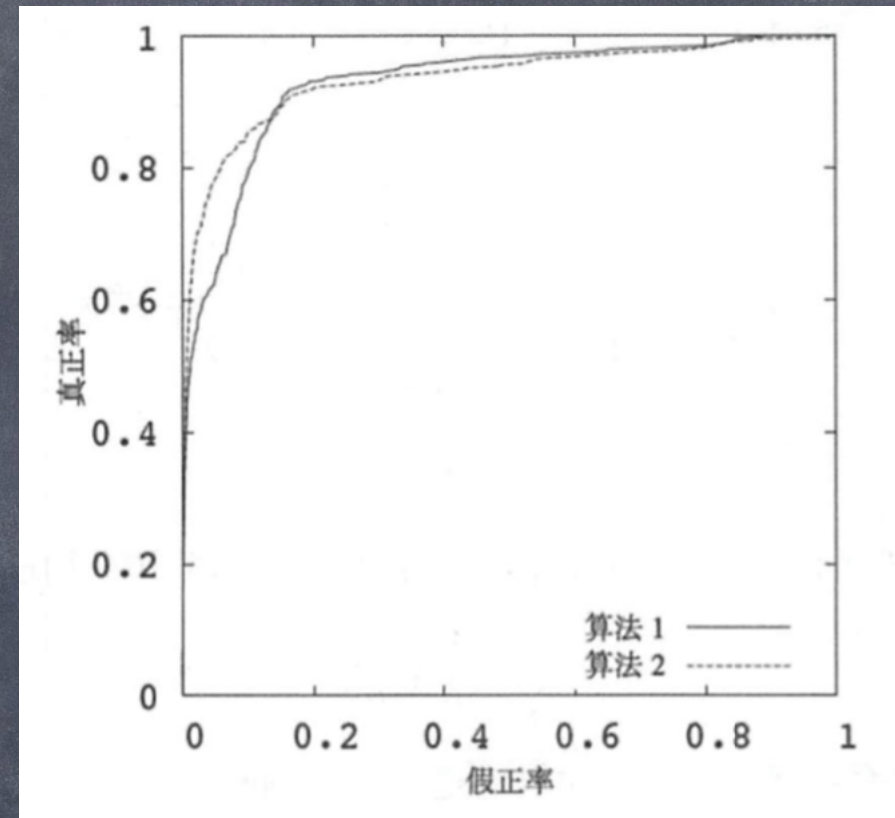
ROC曲线

- 真正率：正样本中预测正确的概率

$$\text{真正率 (TPR)} = \frac{TP}{TP + FN}$$

- 假正率：副样本中预测错误的概率

$$\text{假正率 (FPR)} = \frac{FP}{FP + TN}$$



AUC(对排序敏感)

- 定义式：ROC曲线下面积，积分

$$AUC = \sum_{i \in (P+N)} \frac{(TPR_i + TPR_{i-1}) \cdot (FPR_i - FPR_{i-1})}{2}$$

- 概率式：概率排序，考虑两种极端情况

$$AUC = \frac{\sum_{i \in P} \text{rank}_i - \frac{|P| \cdot (|P| + 1)}{2}}{|P| \cdot |N|}$$

分类指标之对数损失

对数损失(对概率敏感)

二分类损失函数

$$\text{logloss} = -\frac{1}{N} \sum_{i=1}^N (y \log p_i + (y-1) \log (1-p_i))$$

多分类损失函数

$$\text{logloss} = -\frac{1}{N} \cdot \frac{1}{C} \cdot \sum_{i=1}^N \sum_{j=1}^C y_{ij} \cdot \log p_{ij}$$

回归指标

- MAE(Mean Absolute Error)-平均绝对误差, L1范数损失

$$\text{MAE} = \frac{1}{N} \cdot \sum_{i=1}^N |y_i - p_i|$$

- MAPE(Mean Absolute Percentage Error)-平均相对误差

$$\text{MAPE} = \frac{100}{N} \cdot \sum_{i=1}^N \left| \frac{y_i - p_i}{y_i} \right|, y_i \neq 0$$

- RMSE(Root of Mean Squared Error)-均方根误差

$$\text{RMSE} = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^N (y_i - p_i)^2}$$

排序指标

- MAP(Mean Average Precision)-平均准确率均值
- NDCG(Normalized Distributed Cumulative Gain)-归一化累计贴现补贴

排序指标之MAP

- MAP的计算分为两步：

- 计算第K次结果的准确率

$$AP@K = \frac{\sum_{k=1}^{\min(M,K)} P(k) \cdot \text{rel}(k)}{\min(M, K)}$$

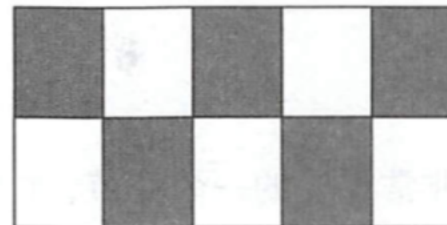
- 计算Q次准确率平均值

$$MAP@K = \sum_{q=1}^Q \frac{AP_q@K}{Q}$$

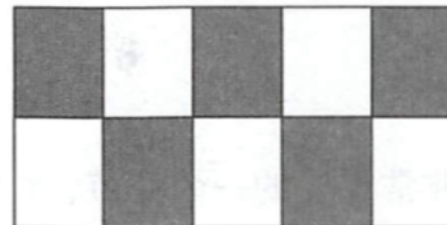
- 举个栗子🌰

下面举个例子说明，其中，黑色代表相关，白色代表不相关。

案例1:



案例2:



- 缺点：无相关权重信息

排序指标之NDCG

• NDCG的计算分3步

— 计算完美排序分数

$$\text{IDCG}@K = \sum_{k=1}^{|\text{REL}|} \frac{2^{\text{rel}_k} - 1}{\log_2(k + 1)}$$

— 计算预测排序分数

$$\text{DCG}@K = \sum_{k=1}^K \frac{2^{\text{rel}_k} - 1}{\log_2(k + 1)}$$

— 预测排序/完美排序

$$\text{NDCG}@K = \frac{\text{DCG}@K}{\text{IDCG}_K}$$

• 举个栗子🍎

文档最佳排序	doc1	doc2	doc3	doc4
最佳排序相应文档相关性 rel_k	2	1	1	0
模型排序	doc3	doc2	doc1	doc4

总结

- AUC vs Logloss
- MAE vs MAPE
- MAE vs RMSE
- MAP vs NDCG