

DM

准确率、精确率、召回率

[【机器学习】准确率、精确率、召回率、误报率、漏报率概念及公式_Better Bench的博客-CSDN博客_漏报率](#)

ROC曲线和AUC值

[机器学习之分类器性能指标之ROC曲线、AUC值 - dzl_ML - 博客园](#)

PR曲线

[准确率、精确率、召回率](#)

lift曲线

[怎么做模型提升度的曲线？](#)

Co-training

[机器学习教程 之 半监督学习 Co-training 协同训练 \(论文、算法、数据集、代码\)_Liangjun_Feng的博客-CSDN博客_协同训练代码（写得太好了）](#)

[协同训练算法之co-training_cccloudcloudy的博客-CSDN博客_co-training（课本翻译版本）](#)

PU learning

[PU-Learning 原理介绍_一个打码的小年轻的的博客-CSDN博客_pu学习](#)

EM算法

[【机器学习】EM——期望最大（非常详细）](#)

[快速讲解](#)

[什么是 EM 算法（最大期望算法）？【知多少】_哔哩哔哩_bilibili](#)

[视频讲解](#)

[数据挖掘十大算法之十——EM算法_哔哩哔哩_bilibili](#)

EM算法 Expectation Maximization → 混合模型 → 聚类

不完备数据

混合: 9 3 8 5 9 2 8 4

算法原理简介

$\theta = (\theta_1, \theta_2)^T$ z_i : 隐变量 $Q(z_i)$: 分布概率

$\sum_{i=1}^n \log P(x_i, \theta) = \sum_{i=1}^n \log \sum_{z_i} P(x_i, z_i, \theta)$

$= \sum_{i=1}^n \log \sum_{z_i} Q(z_i) \cdot \frac{P(x_i, z_i, \theta)}{Q(z_i)}$ $\sum Q(z_i) = 1$

$\geq \sum_{i=1}^n \sum_{z_i} Q(z_i) \log \frac{P(x_i, z_i, \theta)}{Q(z_i)}$ 依下证基高 E步

$\frac{P(x_i, z_i, \theta)}{Q(z_i)} = k \Rightarrow P(x_i, z_i, \theta) = k Q(z_i)$

$\Rightarrow \sum_{z_i} P(x_i, z_i, \theta) = k \sum_{z_i} Q(z_i) = k$

算法步骤:

- (1) 随机初始化模型参数 θ 的初值 θ_0 .
- (2) $j=1, 2, \dots$ 开始EM算法迭代:
 - E步: 计算联合分布的条件概率期望: $Q_i(z_i) = p(z_i | x_i, \theta_j)$
 根据现有模型求分布
 - M步: 极大化 $l(\theta, \theta_j)$, 得到 θ_{j+1} : $l(\theta, \theta_j) = \sum_{i=1}^n \sum_{z_i} Q_i(z_i) \log \frac{p(x_i, z_i, \theta)}{Q_i(z_i)}$
 根据新分布调整模型
- 如果 θ_{j+1} 已经收敛, 则算法结束. 否则继续进行E步和M步进行迭代.

输出: 模型参数 θ .

一个栗子

	Coin A	Coin B
HTTHTHTHTH	0.45 x	0.55 x
HHHHHHHHHH	0.80 x	0.20 x
HTHHHHHHHH	0.73 x	0.27 x
THTTTTTHHT	0.35 x	0.65 x
TTTTTTTTTT	0.65 x	0.35 x

$\theta_A^{(0)} = 0.60$
 $\theta_B^{(0)} = 0.50$

$\theta_A^{(1)} = \frac{21.3}{21.3 + 8.6} = 0.71$
 $\theta_B^{(1)} = \frac{11.7}{11.7 + 8.4} = 0.58$

$\theta_A^{(10)} = 0.80$
 $\theta_B^{(10)} = 0.52$

$Q(z_i) = \frac{P(x_i, z_i, \theta)}{\sum_{z_i} P(x_i, z_i, \theta)} = \frac{P(x_i, z_i, \theta)}{P(x_i, \theta)} = P(z_i | x_i, \theta)$

$a = 0.6^9 \cdot 0.4^1$ $b = 0.5^9 \cdot 0.5^1$ $\frac{a}{a+b} = 0.45$ $\frac{b}{a+b} = 0.55$ Q

3. 使用 EM 的半监督文本分类

朴素贝叶斯分类

二项分布和多项分布 - 小糊涂也学要编程 - 博客园

Naive bayes 朴素简单傻瓜 分类 概率题

一道例题

解答与算法

A	B	C	Tag
A ₁	B ₁	C ₁	T ₁
A ₂	B ₂	C ₂	T ₂
A ₃	B ₂	C ₁	T ₁
A ₁	B ₁	C ₂	T ₂
A ₂	B ₂	C ₂	T ₁
A ₃	B ₃	C ₂	T ₂

A₂B₃C₁ ? 类别 T₁ or T₂

黑客 (A₁B₁C₁) $\frac{P(A_1)P(B_1)P(C_1)}{P(B)}$ $\frac{P(A_1)P(B_1)P(C_1)}{P(B)}$

分类

$P(T_1 | A_2 B_3 C_1) = \frac{P(T_1)P(A_2 B_3 C_1 | T_1)}{P(A_2 B_3 C_1)}$

$P(T_2 | A_2 B_3 C_1) = \frac{P(T_2)P(A_2 B_3 C_1 | T_2)}{P(A_2 B_3 C_1)}$

\downarrow 先验概率 \downarrow 独立性 \downarrow 联合性

$= P(A_2 | T_2) \cdot P(B_3 | T_2) \cdot P(C_1 | T_2)$

\downarrow 统计 \downarrow 数据 \downarrow log 取对数技巧

$1/10^5 \rightarrow$ 采样 数据性材料 $1/10000$ 拉普拉斯平滑

\downarrow 连续性 \downarrow 根号去展 $f(x)$

常见期望和方差

[常见分布 的 数学期望以及方差公式_二喵君的博客-CSDN博客_期望与方差公式汇总](#)

$$D(X)=E\{[X-E(X)]^2\}$$

$$=E\{X^2-2XE(X)+E^2(X)\}$$

因为 $E[-2XE(X)]=-2E^2(X)$ ，所以上式可写成如下：

$$D(X)=E\{X^2-2XE(X)+E^2(X)\}$$

$$=E[X^2-2E^2(X)+E^2(X)]$$

$$=E[X^2-E^2(X)]$$

$$=E(X^2)-E^2(X)$$

常见距离和相似度量

[10种距离和相似性度量（简单介绍）](#)

[常见的9种距离度量方法-电子发烧友网（优缺点）](#)

[全面归纳距离和相似度方法\(7种\)（包括一些信息论的距离）](#)

SVM

[svm原理从头到尾详细推导](#)

[【机器学习】支持向量机 SVM（非常详细）](#)

smo:[快速理解SMO算法_哔哩哔哩_bilibili](#)

kkt条件[KKT条件，原来如此简单 | 理论+算例实践](#)

贝叶斯、极大似然、最大后验

[极大似然估计、最大后验估计](#)

拉普拉斯、高斯分布和L1，L2正则

[Laplace\(拉普拉斯\)先验与L1正则化，Gauss先验导出L2正则化 - Sanny.Liu-CV&&ML - 博客园](#)

[深入理解L1、L2正则化](#)

二项分布和多项式分布

[二项分布和多项分布 - 小糊涂也学要编程 - 博客园](#)

[多项式分布的理解概率公式的理解_猪逻辑公园的博客-CSDN博客_多项式分布](#)

集成学习

[Boosting学习笔记\(Adboost、GBDT、Xgboost\) - Will的笔记 - 博客园](#)

[Bagging和Boosting的总结](#)