



杭州电子科技大学  
HANGZHOU DIANZI UNIVERSITY

篆學易并

育正和薪

CAMALAB  
计算机动画与多媒体分析实验室

# 朴素贝叶斯分类

阮杰

杭州电子科技大学，计算机学院

- 一、对mnist数据集进行朴素贝叶斯分类
- 二、贝叶斯决策论

## 对mnist数据集进行朴素贝叶斯分类

# 1.1 目的

实现这样一个分类器：给定一张图片，分类器给出该数字属于哪个分类（10个分类）

## 1.2 贝叶斯公式

$$P(y^{(i)} = j | x^{(i)}) = \prod_{k=1}^{28*28} \frac{P(x_k^{(i)} | y^{(i)} = j) P(y^{(i)} = j)}{P(x_k^{(i)})}$$

举个例子，从分子上看，若从属于0类的图片在第20个像素上值是1的概率较高（已将图像二值化），那么我们如果发现第20个像素值为1，那么这张图属于0类的概率较高。从分母上看，如果对于所有图片，在第20个像素上1的概率较高，那么就说明这个像素对于分类的帮助较低，因为不管哪一类的图，该像素都可能为1。

# 深入理解公式

$$p(x_k = 1) = \frac{\text{图像中第}k\text{个像素为1的图像数}}{\text{所有图像数}}$$

$$p(y = j) = \frac{\text{属于}j\text{类的图像数}}{\text{所有图像数}}$$

$$p(x_k = 1|y = j) = \frac{\text{属于}j\text{类的图像中像素}k\text{为1的数量}}{\text{属于}j\text{类的图像个数}}$$

# 1.3代码展示

## 贝叶斯决策论



$$P(H|E) = \frac{P(E|H)P(H)}{P(E)}$$

- 贝叶斯决策的基本理论依据就是贝叶斯公式，由总体密度 $P(E)$ 、先验概率 $P(H)$ 和类条件概率 $P(E|H)$ 计算出后验概率 $P(H|E)$ ，判决遵从最大后验概率。

# 证明

以二分类问题为例，对样本 $x$ 的决策错误率如下：

$$P(e|x) = \begin{cases} P(w_2|x) & \text{若决策 } x \in w_1 \\ P(w_1|x) & \text{若决策 } x \in w_2 \end{cases}$$

即

$$P(e|x) = \begin{cases} 1 - P(w_1|x) & \text{若决策 } x \in w_1 \\ 1 - P(w_2|x) & \text{若决策 } x \in w_2 \end{cases}$$

那么最小化错误率的贝叶斯最优分类器为

$$P(e) = \arg \max P(w|x), \quad w \in w_1, w_2$$

# 最小风险贝叶斯决策

除了关心决策正确与否，有时候更得关心错误的决策带来的损失。比如医疗诊断方面，判断细胞是否是癌细胞的决策中，把正常细胞判定为癌细胞，或者把癌细胞判定为正常细胞，这两种决策错误所产生的代价是不一样的。

# 损失函数

设对于实际状态为 $w_j$ 的向量 $x$ 所采取决策 $\alpha_i$ 所带来的损失为

$$\lambda(\alpha_i, w_j), i = 1, \dots, k, j = 1, \dots, c$$

通常损失函数可以用表格的形式给出，叫做决策表。

以判别癌细胞为例，状态1为正常细胞，状态2为癌细胞，假设：

$$\lambda_{11} = 0, \lambda_{12} = 6, \lambda_{21} = 1, \lambda_{22} = 0$$

# 计算步骤

- 利用贝叶斯公式计算后验概率：

$$P(w_j|x) = \frac{p(x|w_j)P(w_j)}{\sum_{i=1}^c p(x|w_i)P(w_i)}, j = 1, \dots, c$$

- 利用决策表，计算条件风险：

$$R(\alpha_i|x) = \sum_{j=1}^c \lambda(\alpha_i|w_j)P(w_j|x), i = 1, \dots, k$$

- 选择风险最小的决策

$$\alpha = \operatorname{argmin}_{i=1, \dots, k} R(\alpha_i|x)$$

# 举例

状态1为正常，状态2为癌细胞，假设：

$$P(w_1) = 0.9, \quad P(w_2) = 0.1$$

$$p(x|w_1) = 0.2 \quad (\text{细胞是正常的诊断错误的概率})$$

$$p(x|w_2) = 0.4 \quad (\text{细胞是癌细胞诊断错误的概率})$$

计算后验概率：

$$P(w_1|x) = 0.818, \quad P(w_2|x) = 0.182$$

计算条件风险：

$$R(\alpha_1|x) = \lambda_{12}P(w_2|x) = 1.092$$

$$R(\alpha_2|x) = \lambda_{21}P(w_1|x) = 0.818$$

# 举例

由于 $R(\alpha_1|x) > R(\alpha_2|x)$ ,即判别为1类的风险更大, 根据最小风险决策, 应该判别为2类, 因为对两类错误带来的风险的认识的不同, 从而产生了与之前不同的决策。

显然, 当对不同类判决的错误风险一致的时候, 最小风险贝叶斯决策就会转化成最小错误率贝叶斯决策。



# THE END

## 感谢观看