# 朴素贝叶斯

W.J.Z

**摘要**：朴素贝叶斯法是基于贝叶斯原理和特征条件独立假设的分类方法。对于给定的输出集，首先基于特征条件独立假设学习输入/输出之间的联合概率分布，然后基于此模型，利用贝叶斯定理求出后验概率最大的输出。

## 0.1 朴素贝叶斯原理

朴素贝叶斯法采用期望风险最小化 (误分类的样本尽可能少)，选择 0-1 损失函数：

$$L(Y,f(X))=\begin{cases}1, & Y\neq f(X)\\ 0, & Y=f(X)\end{cases} \tag{1}$$

期望风险函数为：

$$E\left[L(Y,f(X)\right] \tag{2}$$

注意公式 2是针对的联合分布，而我们需要的是条件期望，根据 0-1 分布期望公式，可得条件期望公式为：

$$\sum_{i=1}^{n}\left(L\left(c_i,f\left(X\right)\right)P\left(c_i|X\right)\right) \tag{3}$$

$c_i\epsilon Y$ 我们要使期望风险最小化，也就是使公式 3最小化，将公式 1带入公式 3得：

$$\begin{aligned}f\left(x\right)&=\arg min\sum_{i=1}^{n}\left(L\left(c_i,f\left(X\right)\right)P\left(c_i|X=x\right)\right)\\&=\arg min\sum_{i=1}^{n}\left(1\times P\left(y\neq c_i|X=x\right)+0\times P\left(y=c_i|X=x\right)\right)\\&=\arg min\sum_{i=1}^{n}\left(P\left(y\neq c_i|X=x\right)\right)\\&=\arg min\left(1-\left(P\left(y=c_i|X=x\right)\right)\right)\\&=\arg maxP\left(y=c_i|X=x\right)\end{aligned}$$

后验概率最大化准则：

$$f\left(x\right)=\arg\underset{c_i}{sum}\left(P\left(Y=c_i|X=x\right)\right) \tag{4}$$

## 0.2 朴素贝叶斯公式

朴素贝叶斯进行了条件独立性假设，条件独立性假设为:

$$P\left(X=x|Y=c_k\right)=P\left(x_1,x_2,\ldots,x_n|Y=c_k\right) \tag{5}$$

$$=\prod_{i=1}^{n}P\left(x_i|Y=c_k\right) \tag{6}$$

后验概率公式:

$$P\left(Y=c_i|X=x\right)=\frac{P\left(X=x|Y=c_i\right)P\left(Y=c_i\right)}{\sum_i P\left(X=x|Y=c_i\right)P\left(Y=c_i\right)} \tag{7}$$

将公式 6带入公式 7得：

$$P\left(Y=c_k|X=x\right)=\frac{P\left(Y=c_k\right)\prod_i P\left(x_i|Y=c_k\right)}{\sum P\left(Y=c_k\right)\prod_i P\left(x_i|Y=c_k\right)} \tag{8}$$

由于分母对于任何 $c_k$ 都是一样的，综上可得朴素贝叶斯分类器可表示为：

$$y=\arg\underset{c_k}{max}P\left(Y=c_k\right)\prod_i P\left(x_i|Y=c_k\right) \tag{9}$$

## 0.3 极大似估计

先验概率 $P(Y=c_k)$ 的极大似然估计为：

$$P\left(Y=c_k\right)=\frac{\sum_{i=1}^{N}I\left(y_i=c_k\right)}{N} \tag{10}$$

条件概率的极大似然估计：

$$P\left(x_i|Y\right)=\frac{\sum_{i=1}^{N}I\left(x_i,y_i=x_k\right)}{\sum_{i=1}^{N}I\left(y_i=c_k\right)} \tag{11}$$

**参考：**统计学习方法