Naive Bayes

$$T=\{(x_1,y_1),(x_2,y_2),\cdots,(x_N,y_N)\} \quad y\in Y=\{c_1,\cdots,c_k\}$$

$$x_i=(x_i^{(1)},x_i^{(2)},\cdots,x_i^{(n)}) \quad 即\ x\in R^n$$

先验概率 $P(Y=c_k)$

先验条件概率 $P(X=x|Y=c_k)=P(X^{(1)}=x^{(1)},\cdots,X^{(n)}=x^{(n)}|Y=c_k)$

$\Downarrow$ 由条件独立假设（"朴素"的原因）

$$=\prod_{j=1}^{n}P(X^{(j)}=x^{(j)}|Y=c_k)$$

后验概率：$P(Y=c_k|X=x)=\dfrac{P(X=x|Y=c_k)P(Y=c_k)}{\sum_{i=1}^{k}P(X=x|Y=c_i)P(Y=c_i)}$

$$=\frac{P(Y=c_k)\cdot\prod_{j=1}^{n}P(X^{(j)}=x^{(j)}|c_k)}{\underbrace{\sum_{i=1}^{k}P(Y=c_i)\cdot\prod_{j=1}^{n}P(X^{(j)}=x^{(j)}|Y=c_i)}_{\text{对于所有}c_k\text{相同}}}=f(x)$$

$\arg\max_{c_k} f(x)$ 等价于 $\arg\max_{c_k}\left(P(Y=c_k)\cdot\prod_{j=1}^{n}P(X^{(j)}=x^{(j)}|Y=c_k)\right)$

模型学习方法：① 极大似然估计 $x^{(j)}$ 可能取值的集合 $\{a_{j1},a_{j2},\cdots,a_{jl}\}$

$$P(X^{(j)}=a_{jl}|Y=c_k)=\frac{\sum_{i=1}^{N}I(x_i^{(j)}=a_{jl},y_i=c_k)}{\sum_{i=1}^{N}I(y_i=c_k)}$$

② 贝叶斯估计

$$P_\lambda(X^{(j)}=a_{jl}|Y=c_k)=\frac{\sum_{i=1}^{N}I(x_i^{(j)}=a_{jl},y_i=c_k)+\lambda}{\sum_{i=1}^{N}I(y_i=c_k)+S_j\lambda}$$