输入	训练数据 $T=\{(x_1,y_1),(x_2,y_2),...,(x_N,y_N)\}$，其中 $x_i=(x_i^{(1)},x_i^{(2)},...,x_i^{(n)})^T$， $x_i^{(j)}$ 是第 i 个样本的第 j 个特征， $x_i^{(1)}\in\{a_{j1},a_{j2},...,a_{jS_j}\}$， a_{jl} 是第 j 个特征可能取的第 l 个值， $j=1,2,\ldots,n$， $l=1,2,\ldots,S_j$， $y_i\in\{c_1,c_2,...,c_K\}$；实例 x。
输出	实例 x 的分类
步骤	（1）计算先验概率及条件概率 极大似然估计 $P(Y=c_k)=\dfrac{\sum_{i=1}^{N}I(y_i=c_k)}{N},k=1,2,...,K$ $P(X^{(j)}=a_{jl}\mid Y=c_k)=\dfrac{\sum_{i=1}^{N}I(x_i^{(j)}=a_{jl},y_i=c_k)}{\sum_{i=1}^{N}I(y_i=c_k)},j=1,2,...,n;l=1,2,...,S_j;k=1,2,...,K$ （2）对于给定的实例 $x=(x^{(1)},x^{(2)},...,x^{(n)})^T$，计算 $P(Y=c_k)\prod_{j=1}^{n}P(X^{(j)}=x^{(j)}\mid Y=c_k),k=1,2,...,K$ （3）确定实例 x 的分类 $y=\arg\max_{c_k}P(Y=c_k)\prod_{j=1}^{n}P(X^{(j)}=x^{(j)}\mid Y=c_k),k=1,2,...,K$
解释	（1）极大似然估计可替换为贝叶斯估计，用来防止极大似然估计会出现估计的概率值为 0 的情况。 贝叶斯估计 $P_\lambda(Y=c_k)=\dfrac{\sum_{i=1}^{N}I(y_i=c_k)+\lambda}{N+K\lambda},k=1,2,...,K$ $P_\lambda(X^{(j)}=a_{jl}\mid Y=c_k)=\dfrac{\sum_{i=1}^{N}I(x_i^{(j)}=a_{jl},y_i=c_k)+\lambda}{\sum_{i=1}^{N}I(y_i=c_k)+S_j\lambda},j=1,2,...,n;l=1,2,...,S_j;k=1,2,...,K$