为了避免其他属性携带的信息被训练集中未出现的属性值"抹去"，在估计概率值时通常要进行"平滑"，常用"拉普拉斯修正"

N表示训练集D中可能的类别数.

$$P(c)=\frac{|D_c|}{|D|}$$

$$\hat{P}(c)=\frac{|D_c|+1}{|D|+N}$$

$$P(x_i|c)=\frac{|D_{c,x_i}|}{|D_c|} \Rightarrow \hat{P}(x_i|c)=\frac{|D_{c,x_i}|+1}{|D_c|+N_i}$$

半朴素贝叶斯分类器.

对"属性条件独立性假设"进行一定程度的放松

III.

EM算法

Z表示隐变量集

Θ表示模型参数

X表示已观测变量.

$$LL(\Theta|X,Z)=\ln P(X,Z|\Theta)$$

$$LL(\Theta|X)=\ln P(X|\Theta)=\ln \sum_Z P(X,Z|\Theta)$$

EM常用来估计参数隐变量，迭代式的方法.

以初始值Θ^0为起点：

E① 基于Θ^t推断隐变量Z的期望，记为Z^t

M② 基于已观测变量X和Z^t对参数Θ做极大似然估计，记为Θ^{t+1}

直到 Θ^t和Z^t稳定

基于Θ^t计算隐变量Z的概率分布$P(Z|X,\Theta^t)$

E① 以当前Θ^t推断$P(Z|X,\Theta^t)$，并计算对数似然

$$LL(\Theta|X,Z)$$

$$Q(\Theta|\Theta^t)=E_{Z|X,\Theta^t}LL(\Theta|X,Z)$$