what

Q How

A.

1. 得到先验概率分布

$$P(Y=C_k) \ ; \quad k=1,2,\cdots k.$$

2. 条件概率分布

$$P(X=x|Y=C_k)=P(X^{(1)}=x^{(1)},\cdots,X^{(n)}=x^{(n)}|Y=C_k)$$

3. 学习到联合概率分布 $P(X,Y)$

- 朴素贝叶斯法对条件概率分布作了条件独立性假设.

$$P(X=x|Y=C_k)=P(X^{(1)}=x^{(1)},\cdots,X^{(n)}=x^{(n)}|Y=C_k)$$

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

- 朴素贝叶斯法实际上学习到生成数据的机制，属于生成模型.

Q 分类

A.

$$P(Y=C_k|X=x)=\frac{P(X=x|Y=C_k)P(Y=C_k)}{\sum_k P(X=x|Y=C_k)P(Y=C_k)}$$

条件独立性,

$$P(Y=C_k|X=x)=\frac{P(Y=C_k)\prod P(X^{(j)}=x^{(j)}|Y=C_k)}{\sum P(Y=C_k)\prod P(X^{(j)}=x^{(j)}|Y=C_k)}$$

$$y=\arg\max_{C_k} P(Y=C_k|X=x)$$

$$y=\arg\max_{C_k} P(Y=C_k)\prod P(X^{(j)}=x^{(j)}|Y=C_k).$$

# why 朴素贝叶斯法很 cool

- inference cheap
- 参数量少
- 利用标签数据估计参数
- 联系分类与回归
- 实践成功