

7.1 贝叶斯决策论

a. 如何基于这些根先验和误判损失来选择最优的类别标记。

$$R(G|x) = \sum_{j=1}^N \lambda_{ij} P(G|c) \quad \begin{array}{l} \text{条件风险} \\ \text{C 被判为类的误判损失} \end{array} \quad \begin{array}{l} \text{在已知根先验} \\ \text{x 分类为 G 的期望损失} \end{array}$$

b. 贝叶斯判定准则: 为最小化总条件风险, 只需在每个样本上选择 $R(c|x)$ 最小。

$$h^*(x) = \arg \min_{c \in Y} R(c|x) \quad \text{最优分类器}$$

若 $\lambda_{ij} = \begin{cases} 0 & i=j \\ 1 & i \neq j \end{cases}$ 则 $R(c|x) = 1 - P(c|x)$

c. 求 $P(c|x)$ $\begin{cases} \text{i. 给定 } x, \text{ 建模 } P(c|x) \text{ 来预测 } c. & (\text{判别式 model}) \\ \text{ii. 对联合分布 } P(x, c) \text{ 建模, } \Rightarrow P(c|x) & (\text{生成式 model}) \end{cases}$

$$P(c|x) = \frac{P(c)P(x|c)}{P(x)}$$

$$P(c) = \frac{|D_c|}{|D|} \rightarrow \text{属于 } c \text{ 的样本数} / \text{总样本数}$$









