

朴素贝叶斯

条件独立性假设: $P(x|y) = \prod_{i=1}^p P(x_i|y)$ 向后遍算

Data: $\{(x_i, y_i)\}_{i=1}^N \quad x_i \in \mathbb{R}^p \quad y_i \in \{0, 1\}$

给定 $x \rightarrow y$ 是 y_0 ?

$$y = \arg \max_y P(y|x) = \arg \max_{y \in \{0, 1\}} \frac{P(x, y)}{P(x)}$$

$$P(y|x) = \frac{P(x, y)}{P(x)} = \frac{P(y) P(x|y)}{P(x)} \propto P(y) P(x|y)$$

$$= \arg \max_y P(y) \cdot P(x|y)$$

二分类: 0/1 \rightarrow 伯努利分布

多分类: 类别分布

$$P(x_i|y) \left\{ \begin{array}{l} \text{离散} \rightarrow x_i \Rightarrow \text{Categorical distribution} \\ \text{连续} \rightarrow x_i \Rightarrow N(\mu_i, \sigma_i^2) \end{array} \right.$$