Naive Bayes Classifier

最简单的有向概率图

$\hat{y} = \arg\max_{y} P(y|x)$

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

$= \arg\max_{y} P(y) \cdot P(x|y)$

条件独立性假设：$P(x|y) = \prod_{j=1}^{p} P(x_j|y)$

$\begin{cases} x\text{离散} \rightarrow x_j \sim \text{Categorical Dist} \\ x\text{连续} \rightarrow x_j \sim N(\mu_j, \sigma_j^2) \end{cases}$

(五) 降维．(协方差矩阵计算的协方差是feature之间的协方差)．

Vazyme

过拟合 $\begin{cases} \text{data} \uparrow \\ \text{正则化} \\ \text{降维} \end{cases}$

降维 $\rightarrow$ $\begin{cases} \text{直接降维 (特征选择) [lasso]} \\ \text{线性降维} \\ \text{非线性降维} \end{cases}$

Sample Mean: $\bar{X}_{p\times 1} = \frac{1}{N}\sum_{i=1}^{N} x_i = \frac{1}{N}(x_1 \cdots x_N)\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}_{N\times 1} = \frac{1}{N}X^T 1_N$

Sample Covariance: $S_{p\times p} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})(x_i - \bar{x})^T$

$S = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})(x_i - \bar{x})^T \quad = \frac{1}{N}X^T H X$

$= \frac{1}{N}\sum_{i=1}^{N}(x_i - \frac{1}{N}X^T 1_N)(x_i - \frac{1}{N}X^T 1_N)^T$

$= \frac{1}{N}\underbrace{(x_1 - \bar{x}, x_2 - \bar{x}, \cdots, x_N - \bar{x})}_{(x_1, x_2, \cdots x_N) - (\bar{x}, \bar{x}, \cdots \bar{x})}\begin{pmatrix}(x_1-\bar{x})^T \\ (x_2-\bar{x})^T \\ \vdots \\ (x_N-\bar{x})^T\end{pmatrix}$

$= X^T - \bar{x}\,(1, 1, \cdots 1) = X^T - \bar{x}1_N^T = X^T - \frac{1}{N}X^T 1_N 1_N^T$

$= \frac{1}{N}X^T \underbrace{(I_N - \frac{1}{N}1_N 1_N^T)}_{H_N \text{ centering matrix}} \cdot (I_N - \frac{1}{N}1_N 1_N^T)^T X$

$= \frac{1}{N}X^T H \cdot H^T X \Rightarrow \frac{1}{N}X^T H X$

$H^T = H$

$H^2 = H \cdot H = (I_N - \frac{1}{N}1_N 1_N^T)(I_N - \frac{1}{N}1_N 1_N^T)$

$= I - \frac{2}{N}1_N 1_N^T + \frac{1}{N^2}1_N 1_N^T 1_N 1_N^T$