Sep 3, 2024

Bayes Rule

$$P(Y|x) = \frac{P(x, Y)}{P(X)} = \frac{P(X|y)\cdot P(y)}{\sum_{y'} P(x|y')\cdot P(y')}$$

$$P(y|x) \propto P(x|y)\cdot P(y)$$

$$\underline{D = \{1, 0, 1, 1\}}$$

$X \sim Ber(\theta) \quad 0 \leq \theta \leq 1$ θ ?

likelihood $P(D|\theta)$

$$= \theta\cdot(1-\theta)\cdot\theta\cdot\theta$$

$$= \theta^3\cdot(1-\theta)$$

$D = \{N_1$ positive samples, N_0, negative samples$\}$

likelihood $P(D|\theta)$

$$= \theta^{N_1}(1-\theta)^{N_0}$$

θ? $\boxed{\text{MLE. maximum likelihood Estimate}}$

$$\hat{\theta}_{MLE} = \arg\max_{\theta} \underline{P(D|\theta)}$$

$$\frac{\partial P(D|\theta)}{\partial\theta} = 0$$

$$\frac{\partial\, \theta^{N_1}(1-\theta)^{N_0}}{\partial \theta} = 0$$

$$N_1 \theta^{N_1-1}(1-\theta)^{N_0} - \theta^{N_1} N_0 (1-\theta)^{N_0-1} = 0$$

$$\theta = \frac{N_1}{N_1+N_0} = \frac{N_1}{N} \qquad N = N_1 + N_0$$

posterior $\propto$ likelihood $\times$ piror

$$P(\theta \mid D) \;\underline{\underline{\propto}}\; P(D \mid \theta) \cdot P(\theta)$$

$$p(\theta \mid D) \propto \theta^{N_1+a-1}(1-\theta)^{N_0+b-1}$$

$$\hat{\theta}_{MAP} = \underset{\theta}{\operatorname{argmax}}\; p(\theta \mid D)$$

$$\boxed{= \frac{N_1 + a - 1}{N + a + b - 2} \qquad \text{MAP}}$$

Bayesian Averaging

$$P(x^* = 1 \mid D) = \int_0^1 p(x^* = 1 \mid \theta)\, p(\theta \mid D)\, d\theta$$

$$= \int_0^1 \theta \cdot \mathrm{Beta}(\theta \mid N_1 + a, N_0 + b)\, d\theta$$

$$= \frac{N_1 + a}{N + a + b}$$