3 - BAYESIAN LEARNING

Aprendizagem 2024/2025

TEOREMA DE BAYES

- As crenças (A) podem ser suportadas por evidências (B), dada pelo seguinte teorema:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

- O valor da previsão $\hat{y}$ é dado por:

$$\hat{y} = \underset{c_i}{\operatorname{argmax}}\{p(c_i|x)\} = \underset{c_i}{\operatorname{argmax}}\left\{\frac{p(x|c_i)p(c_i)}{p(x)}\right\}$$

ASSUMPTIONS

- **Maximum a posteriori** (o valor do denominador vai ser igual para todos os valores calculados):

$$\hat{y}_{MAP} = \underset{c_i}{\operatorname{argmax}}\{p(c_i|x)\} = \underset{c_i}{\operatorname{argmax}}\left\{\frac{p(x|c_i)p(c_i)}{\cancel{p(x)}}\right\}$$

- **Maximum likelihood** (se assumirmos que todas as crenças têm a mesma probabilidade):

$$\hat{y}_{ML} = \underset{c_i}{\operatorname{argmax}}\{p(c_i|x)\} = \underset{c_i}{\operatorname{argmax}}\left\{\frac{p(x|c_i)\cancel{p(c_i)}}{\cancel{p(x)}}\right\}$$

NAÏVE BAYES

- As variáveis são condicionalmente independentes, e dado pela seguinte fórmula:

$$\hat{y}_{MAP} = \underset{c_i}{\operatorname{argmax}}\{p(c_i|x)\} = \underset{c_i}{\operatorname{argmax}}\left\{p(c_i)\prod_{j=1}^{d} p(x_j|c_i)\right\}$$

CÁLCULO DAS PROBABILIDADES

- Variável **categórica**: cálculo direto da probabilidade
- Variável **numérica**: pode seguir uma distribuição normal

$$p(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

- Variáveis **numéricas**: pode seguir uma distribuição normal multivariada

$$p(x|\mu,\Sigma) = \frac{1}{(2\pi)^{\frac{m}{2}}\sqrt{|\Sigma|}} e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)}$$

OUTRAS FÓRMULAS

- **Matriz de covariâncias** entre y1 e y2 (Σ):

$$\begin{pmatrix} {\sigma_{y_1}}^2 & cov(y2, y1) \\ cov(y1, y2) & {\sigma_{y_2}}^2 \end{pmatrix}$$

$${\sigma_{y_1}}^2 = \frac{1}{n-1} \sum (y_{1_i} - \mu_1)^2$$

$$cov(y1, y2) = cov(y2, y1) = \frac{1}{n-1} \sum (y_{1_i} - \mu_1)(y_{2_i} - \mu_2)$$

SUMÁRIO

- Ficha 3