Exercise 16

Freitag, 10. Juni 2022 12:21

a)

$P(S \cap W) = P(W \cap S)$ I

$P(S \cap W) = P(S)\,P(W|S)$ II

$P(W \cap S) = P(W)\,P(S|W)$ III

I states that II = III

$\Rightarrow P(W)\,P(S|W) = P(S)\,P(W|S)$

$\Leftrightarrow P(S|W) = \dfrac{P(W|S) \cdot P(S)}{P(W)}$

b)

$P(S) = \dfrac{\text{Times soccer was played}}{\text{was played + was not played}} = \dfrac{9}{14}$

$P(W|S) = \prod_i P(x_i|S)$ $\qquad P(x_i|S) = P(S \cap x_i) : P(S)$

$P(\text{wind} = \text{high} \mid S = \text{yes}) = \dfrac{P(S = \text{yes}, \text{wind} = \text{high})}{P(S)} = \dfrac{3}{14} \cdot \dfrac{14}{9} = \dfrac{1}{3}$

$P(\text{humidity} = \text{high} \mid S = \text{yes}) = \dfrac{3}{9} = \dfrac{1}{3}$ (Table 1: row 2, column 1: 3 times high humidity when soccer was played (9 times total))

$P(\text{temperature} = \text{cold} \mid S = \text{yes}) = \dfrac{1}{3}$

$P(\text{forecast} = \text{sunny} \mid S = \text{yes}) = \dfrac{2}{9}$

$\Rightarrow P(W|S) = \left(\dfrac{1}{3}\right)^3 \cdot \dfrac{2}{9} = \dfrac{2}{243}$

$P(W = \text{table 2}) = P(\text{wind} = \text{high}) \cdot P(\text{humidity} = \text{low}) \cdot P(\text{temperature} = \text{cold}) \cdot P(\text{forecast} = \text{sunny})$

$= \dfrac{6}{14} \cdot \dfrac{7}{14} \cdot \dfrac{6}{14} \cdot \dfrac{3}{14}$

$= \dfrac{27}{1372}$

$\Rightarrow P(S = \text{yes} \mid W = \text{table 2}) = \dfrac{2}{243} \cdot \dfrac{9}{14} \cdot \dfrac{1372}{27} = \dfrac{196}{729}$

$\approx 26{,}89\,\%$

c)

The problem that occurs is, that there is no data for hot weather when soccer was played. That leads to $P(\text{temperature} = \text{hot} \mid S = \text{yes}) = 0$

$\Rightarrow P(W = \text{table 3} \mid S = \text{yes}) = 0$ even though the propability should be > 0

To solve that we can calculate $P(S = \text{yes} \mid W) = 1 - P(S = \text{no} \mid W)$

$P(S = \text{no} \mid W) = \dfrac{P(W \mid S = \text{no}) \cdot P(S = \text{no})}{P(W)}$

$P(S = \text{no}) = \dfrac{5}{14}$

$P(W \mid S = \text{no}) = \dfrac{2}{5} \cdot \dfrac{4}{5} \cdot \dfrac{1}{5} \cdot \dfrac{1}{5} = \dfrac{8}{625}$

$P(W) = \dfrac{8}{14} \cdot \dfrac{7}{14} \cdot \dfrac{1}{14} \cdot \dfrac{3}{14} = \dfrac{3}{686}$

$\Rightarrow P(S = \text{yes} \mid W = \text{table 3}) = 1 - \dfrac{8}{625} \cdot \dfrac{5}{14} \cdot \dfrac{686}{3}$

Exercise 17

a) Y: soccer

$$H(Y) = -P(Y=true)\log_2(P(Y=true)) - P(Y=false)\log_2(P(Y=false))$$

$$= -\frac{9}{14}\log_2\left(\frac{9}{14}\right) - \frac{5}{14}\log_2\left(\frac{5}{14}\right)$$

$$\approx \underline{\underline{0{,}94}}$$

b) X: wind

$$H(Y|X) = P(X=true)\cdot H(Y|X=true) + P(X=false)\cdot H(Y|X=false)$$

$$H(Y|X=true) = -\frac{2}{6}\log_2\left(\frac{2}{6}\right) - \frac{4}{6}\log_2\left(\frac{4}{6}\right)$$

$$\approx 0{,}918$$

$$H(Y|X=false) = -\frac{6}{8}\log_2\left(\frac{6}{8}\right) - \frac{2}{8}\log_2\left(\frac{2}{8}\right)$$

$$\approx 0{,}811$$

$$\Rightarrow H(Y|X) = \frac{6}{14}\cdot H(Y|X=true) + \frac{8}{14}H(Y|X=false)$$

$$= \underline{\frac{6}{7}}$$

$$\Rightarrow IG(X,Y) = H(Y) - H(Y|X)$$

$$= 0{,}94 - \frac{6}{7} \approx \underline{\underline{0{,}08314}}$$