Sheesh 8

Dienstag, 14. Juni 2022 18:14

16)

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

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

b) Total possible games: 14, played = 9, not = 5

Wind = high – 3x played
3x not

Humidity = high – 3x played
4x not

Temperature = cold – 3x played
3x not

Forecast = sunny – 2x played
1x not

$$P(W) = P(W|played)\, P(played) + P(W|not)\, P(not)$$

$P(high|played) = \frac{1}{3}$, $P(high|played) = \frac{1}{3}$
$P(sunny|played) = \frac{2}{3}$, $P(cold|played) = \frac{1}{3}$

$\Rightarrow P(W|played) = \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{3^5}$

$P(played) = \frac{9}{14}$

$P(high|not) = \frac{3}{5}$; $P(high|not) = \frac{4}{5}$

$P(sunny|not) = \frac{1}{5}$; $P(cold|not) = \frac{3}{5}$

$P(not) = \frac{5}{14}$; $P(W) = \frac{611}{23625}$

$$P(played|W) = \frac{P(W|played) \cdot P(played)}{P(W)} \approx 20{,}46\%$$

c) The Problem is that hot Temperature never occured in this 14 possible games so the Proporbility for playing gets 0 since $P(W|played) = 0$.

$P(low|not) = \frac{2}{5}$; $P(high|not) = \frac{4}{5}$

$P(hot|not) = \frac{1}{5}$; $P(sunny|not) = \frac{1}{5}$

$P(W|not) = \frac{2}{5} \cdot \frac{4}{5} \cdot \frac{1}{5} \cdot \frac{1}{5} = \frac{8}{5^4}$

$P(not) = \frac{5}{14}$

3

$P(w|not) = \frac{2}{5} \cdot \frac{4}{5} \cdot \frac{1}{5} \cdot \frac{1}{5} = \frac{8}{5^4}$

$P(not) = \frac{5}{14}$

$P(low|played) = \frac{6}{14}$; $P(high|played) = \frac{3}{14}$

$P(hot|played) = 0$; $P(sunny|played) = \frac{2}{14}$

2(+1) ideas how to solve:

1. $P(w) = P(w_1) \cdot P(w_2)\, P(w_3)\, P(w_4)$

$= P(low) \cdot P(high) \cdot P(hot) \cdot P(sunny)$

$= \frac{4}{7} \cdot \frac{1}{2} \cdot \frac{1}{14} \cdot \frac{3}{14} = \frac{3}{686}$

$\Rightarrow P(not|w) = \frac{\frac{8}{625} \cdot \frac{5}{14}}{\frac{3}{686}} \approx 0{,}976$

$\Rightarrow 1 - P(not|w) = 2{,}4\%$

2. not consider the temperature:

$$P(not|w) = \frac{\frac{8}{5^3} \cdot \frac{1}{5}}{\frac{8}{5^4} + \frac{36}{14^4} \cdot 9} = 60{,}28\%$$

$\Rightarrow P(played|w) = 1 - P(not|w) = 39{,}72\%$

(3. more data!!!)

17)

a) $H(Y) = -\sum_{z \in Z} P(Y=z) \log_2 P(Y=z)$

$= -\left(\frac{n_{true}}{n} \log_2 \frac{n_{True}}{n} + \frac{n_{false}}{n} \log_2 \frac{n_{false}}{n} \right)$

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

$= 0{,}940286$

b)

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

$H(Y|wind) = -\sum_{g \in \{True, false\}} P(X_{wind} = g)\, H(Y|X=g)$

$n_{soccer,wind}$

$H(Y|X_{wind} = True) = -\left(\frac{n_{True,true}}{n_{True}} \log_2 \frac{n_{True,true}}{n_{True}} + \frac{n_{false,true}}{n_{true}} \log_2 \frac{n_{false,true}}{n_{true}} \right)$

n_{wind}

$= -\left(\frac{3}{6} \log_2 \frac{1}{2} + \frac{1}{2} \log_2 \frac{1}{2} \right) = 1$

$H(Y|W_{wind} = false) = -\frac{6}{8} \log_2 \frac{6}{8} - \frac{2}{8} \log_2 \frac{2}{8} = 0{,}8112781$

$IG(X,Y) = H(Y) - \frac{n_{True}}{n} H(Y|X_{wind} = True) - \frac{n_{false}}{n} H(Y|X_{wind} = false)$

$= 0{,}048127$