Exercises Naive Bayes

1 Naive Bayes

Given is the following data set with four features and one target (*Class*).

$P(A=1 \mid Y=1) = \frac{4}{7}$

$P(A=0 \mid Y=1) = \frac{3}{7}$

$P(A=1 \mid Y=0) = \frac{3}{5}$

$P(A=0 \mid Y=0) = \frac{2}{5}$

$P(B=1 \mid Y=1) = \frac{2}{7}$

$P(B=0 \mid Y=1) = \frac{5}{7}$

$P(B=1 \mid Y=0) = \frac{5}{5} = 1$

$P(B=0 \mid Y=0) = \frac{0}{5} = 0$

$\text{mean}(C \mid Y=1) = \frac{69}{7} = 9{,}857 = \mu$

$\text{var}(C \mid Y=1) = 10{,}476 = \sigma^2$

$\rightarrow P(C \mid Y=1) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

$\text{mean}(C \mid Y=0) = 6{,}4 = \mu_2$

$\text{var}(C \mid Y=0) = 5{,}3 = \sigma_2^2$

$\rightarrow P(C \mid Y=0) = \frac{1}{\sqrt{2\pi\sigma_2^2}} \cdot e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}}$

$P(D=1 \mid Y=1) = \frac{4}{7}$

$P(D=0 \mid Y=1) = \frac{3}{7}$

$P(D=1 \mid Y=0) = \frac{3}{5}$

$P(D=0 \mid Y=0) = \frac{2}{5}$

A	B	C	D	Class (Y)
0	0	10	1	1
0	0	11	0	1
0	1	9	1	0
0	1	8	0	1
0	1	4	0	0
1	0	5	1	1
1	0	12	1	1
1	0	15	0	1
1	1	8	1	0
1	1	7	0	0
1	1	8	1	1
1	1	4	1	0

= 7

= 5

$P(Y=1) = \frac{7}{12}$

$P(Y=0) = \frac{5}{12}$

1. Compute the conditional probabilities $P(X \mid Class)$ where $X \in \{A, B, C, D\}$. Use a normal distribution for C.
2. Classify the instance using Naive Bayes $(1, 1, 5, 0)$. $= \underline{X}$

$\rightarrow P(Y=0 \mid \underline{X}) = P(\underline{X} \mid Y=0) \cdot P(Y=0) = 0{,}0162$

$P(Y=1 \mid \underline{X}) = P(\underline{X} \mid Y=1) \cdot P(Y=1) = 0{,}002039$

$$P(\underline{X} \mid Y=0) = \underbrace{\frac{3}{5}}_{P(A=1 \mid Y=0)} \cdot \underbrace{1}_{P(B=1 \mid Y=0)} \cdot \underbrace{0{,}162}_{ef(5;\mu_2,\sigma_2)} \cdot \underbrace{\frac{2}{5}}_{P(D=0 \mid Y=0)} = 0{,}03888$$

$$ef(5;\mu_2,\sigma_2) = \int_5^{5+1} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}}\, dx \approx 0{,}162$$

mit $\mu_2 = 6{,}4$ und $\sigma_2^2 = 5{,}3$

$$P(\underline{X} \mid Y=1) = \underbrace{\frac{4}{7}}_{P(A=1 \mid Y=1)} \cdot \underbrace{\frac{2}{7}}_{P(B=1 \mid Y=1)} \cdot \underbrace{0{,}04997}_{ef(5;\mu,\sigma)} \cdot \underbrace{\frac{3}{7}}_{P(D=0 \mid Y=1)} = 0{,}003496$$

$$ef(5;\mu,\sigma) = \int_5^{5+1} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx \approx 0{,}04997$$

mit $\mu = 9{,}857$ und $\sigma^2 = 10{,}476$

$\Downarrow$

$$\boxed{Y=0 \rightarrow \text{Class } 0 \text{ am warscheinlichsten}}$$