

Naive Bayes with numeric features;

INPUT

temperature	play golf?
85	no
80	no
83	yes
70	yes
68	yes
65	no
64	yes
72	no
69	yes
75	yes
75	yes
72	yes
81	yes
71	no

QUESTION:

Predict play golf for temperature = 66.

$$\begin{aligned} * P(\text{play}=\text{yes} | \text{temperature}=66) \\ = \frac{P(\text{play}=\text{yes}) \cdot P(\text{temperature}=66 | \text{play}=\text{yes})}{P(\text{temperature}=66)} \end{aligned}$$

$$\begin{aligned} * P(\text{play}=\text{no} | \text{temperature}=66) \\ = \frac{P(\text{play}=\text{no}) \cdot P(\text{temperature}=66 | \text{play}=\text{no})}{P(\text{temperature}=66)} \end{aligned}$$

$$P(\text{play}=\text{yes}) = \frac{9}{14}, \quad P(\text{play}=\text{no}) = \frac{5}{14}$$

How to get $P(\text{temp}=66 | \text{play}=\text{yes})$ and $P(\text{temp}=66 | \text{play}=\text{no})$? Assume $P(\text{temp}=x | \text{play}=\text{yes})$ and $P(\text{temp}=x | \text{play}=\text{no})$ are normally distributed with μ and σ estimated from the training data set.

How to get the σ and μ ?

For play = yes, the temp. values are: 83, 70, 68, 64, 69, 75, 75, 72, 81

$$\text{Their mean: } \mu = \frac{1}{n} \sum x_i = 73$$

$$\begin{aligned} \text{Their } \sigma &= \sqrt{\frac{1}{n-1} \sum (x_i - \mu)^2} = \sqrt{\frac{1}{8} [(83-73)^2 + (70-73)^2 + \dots + (81-73)^2]} \\ &= \sqrt{\frac{1}{8} \cdot 304} = 6.2 \end{aligned}$$

For play = no, the temp. values are: 85, 80, 65, 72, 71

$$\mu = 74.6, \quad \sigma = 7.9$$

Recall the probability density function for a normal variable:

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\begin{aligned} \text{So, } P(\text{temp}=66 | \text{play}=\text{yes}) &= f(66) \text{ for } \mu=73 \text{ and } \sigma=6.2 \\ &= \frac{1}{\sqrt{2\pi} \cdot 6.2} e^{-\frac{(66-73)^2}{2 \cdot 6.2^2}} = 0.034 \end{aligned}$$

And for $P(\text{temp} = 66 | \text{play} = \text{no}) = f(66)$ for $\mu = 74.6$ and $\sigma = 7.9$

$$= \frac{1}{\sqrt{2\pi} \cdot 7.9} e^{-\frac{(66-74.6)^2}{2 \cdot 7.9^2}} = 0.022.$$

So, $P(\text{play} = \text{yes} | \text{temp} = 66) = \frac{P(\text{play} = \text{yes}) \cdot P(\text{temp} = 66 | \text{play} = \text{yes})}{P(\text{temp} = 66 | \text{play} = \text{yes}) \cdot P(\text{play} = \text{yes}) + P(\text{temp} = 66 | \text{play} = \text{no}) \cdot P(\text{play} = \text{no})}$

$$= \frac{\frac{9}{14} \cdot 0.034}{\frac{9}{14} \cdot 0.034 + \frac{5}{14} \cdot 0.022}$$

$$= \frac{0.021857}{0.021857 + 0.007857} = 0.74$$

Since play is a binary class variable,

$$P(\text{play} = \text{no} | \text{temp} = 66) = 1 - 0.74 = 0.26.$$

~~So we~~

So we Predict $\text{play} = \text{yes}$ given $\text{temp} = 66$.