Machine Learning: Homework 3

Laurent HAYEZ

October 20, 2016

Exercise 1. *Invent the two characters Taylor and Robin each with a weight in the range of [63-74] kg and a shoe size in the range of [40-44]. Decide (by hand) with Naïve Bayes using the probability density function whether your Taylor and Robin are female or male based on the following statistics.*

Solution. As we have to deal with numerical attributes, we need to use a probability density function

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

where μ is the sample mean and σ is the standard deviation.

In Table 1, we represented the weight and shoe size according to whether the person is male or female, and we computed the sample means and standard deviations.

Let f_1 be a gaussian distribution with $\mu = \mu_1$ and $\sigma = \sigma_1$ representing the distribution of the weight for the males. Let g_1 be the same but with $\mu = \mu_2$ and $\sigma = \sigma_2$ representing the distribution of the weight for the females. In the same fashion we define f_2 and g_2 corresponding to the shoe size. Moreover let $L[\cdot]$ denote the likelihood of an event.

If Taylor weights 73 kgs and has shoe size 41, and Robin weights 67 kgs and has shoe size 44, then define $E_1 = \{\text{weight} = 73, \text{shoe size} = 41\}$ and $E_2 = \{\text{weight} = 67, \text{shoe size} = 44\}$. We have

$$\begin{aligned} L[\text{Taylor} = \text{male} \mid E_1] &= f_1(73) \cdot f_2(41) \cdot \frac{1}{2} \qquad \text{where } \mathbb{P}[\text{male}] = \frac{1}{2} \\ &= 0.0505 \cdot 0.0118 \cdot \frac{1}{2} \\ &= 0.00029 \end{aligned}$$

□