CS446: Machine Learning, Fall 2017, Homework 1

Name: Lanxiao Bai (lbai5)

Worked individually

Problem 1

Solution: By Bayes Rule, we have that

$$P(y = 1 \mid \mathbf{x}) = \frac{P(\mathbf{x} \mid y = 1)P(y = 1)}{P(\mathbf{x})}$$

Problem 2

Solution: Assume there is a relation a that

$$\log \frac{P(\mathbf{x} \mid y = 1)}{P(\mathbf{x} \mid y = 0)} = a$$

Then we apply Bayes Rule

$$\begin{aligned} &\log \frac{P(\mathbf{x} \mid y = 1)}{P(\mathbf{x} \mid y = 0)} = a \\ \Rightarrow &\log \frac{P(y = 1 \mid \mathbf{x})}{P(y = 0 \mid \mathbf{x})} + \log \frac{P(y = 1)}{P(y = 0)} = a \end{aligned}$$

Let

$$a' = a - \log \frac{P(y = 1)}{P(y = 0)}$$

then we have

$$\begin{aligned} &\log \frac{P(y = 1 \mid \mathbf{x})}{P(y = 0 \mid \mathbf{x})} = a' \qquad (1) \\ \Rightarrow &\frac{P(y = 1 \mid \mathbf{x})}{1 - P(y = 1 \mid \mathbf{x})} = e^{a'} \\ \Rightarrow &\frac{1 - P(y = 1 \mid \mathbf{x})}{P(y = 1 \mid \mathbf{x})} = e^{-a'} \\ \Rightarrow &1 - P(y = 1 \mid \mathbf{x}) = e^{-a'} P(y = 1 \mid \mathbf{x}) \\ \Rightarrow &P(y = 1 \mid \mathbf{x}) = \frac{1}{1 + e^{-a'}} \end{aligned}$$

Problem 3

Solution: Since $\mathbf{x} \sim \mathcal{N}(\mu_c, \mathbf{\Sigma})$, and the result of previous questions, we have

$$\begin{aligned} P(\mathbf{x} \mid y = c) &= \prod_{i=1}^{n} P(x_i \mid y = c) \\ &= \prod_{i=1}^{n} \mathcal{N}(x_i \mid \mu_{ic}, \sigma_i^2) \\ &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x_i-\mu_{ic})^2}{2\sigma_i^2}} \\ &= \left(\frac{1}{\sqrt{2\pi}}\right)^n \frac{1}{\prod_{i=1}^{n} \sigma_i^2} e^{-\sum_{i=0}^{n} \frac{(x_i-\mu_{ic})^2}{2\sigma_i}} \end{aligned} \tag{2}$$

Problem 4

Solution: Since by Bayes rule we have

$$\begin{aligned} P(y = 1 \mid \mathbf{x}) &= \frac{P(y = 1)P(\mathbf{x} \mid y = 1)}{P(y = 1)P(\mathbf{x} \mid y = 1)P(y = 0)P(\mathbf{x} \mid y = 0)} \\ &= \frac{1}{1 + \frac{P(y=0)P(\mathbf{x}|y=0)}{P(y=1)P(\mathbf{x}|y=1)}} \\ &= \frac{1}{1 + \exp(\log \frac{P(y=0)P(\mathbf{x}|y=0)}{P(y=1)P(\mathbf{x}|y=1)})} \\ &= \frac{1}{1 + \exp(\log \frac{P(y=0)P(\mathbf{x}|y=0)}{P(y=1)P(\mathbf{x}|y=1)})} \end{aligned}$$

Since the independence assumption of Naive Bayes, we have

$$\begin{aligned} P(y = 1 \mid \mathbf{x}) &= \frac{1}{1 + \exp(\log \frac{P(y=0)}{P(y=1)} + \log \frac{P(\mathbf{x}|y=0)}{P(\mathbf{x}|y=1)})} \\ &= \frac{1}{1 + \exp(\frac{1-\pi}{\pi} + \sum_{i=1}^{n} \log \frac{P(x_i|y=0)}{P(x_i|y=1)})} \end{aligned} \tag{3}$$

in which

$$\sum_{i=1}^{n} \log \frac{P(x_i \mid y = 0)}{P(x_i \mid y = 1)} = \sum_{i=1}^{n} \left(\frac{\mu_{i0} - \mu_{i1}}{\sigma_i^2} x_i + \frac{\mu_{i0}^2 - \mu_{i1}^2}{\sigma_i^2} \right) (\textit{Mitchell}\ (1997))$$

So that when we let $w_i = \frac{\mu_{i0}-\mu_{i1}}{\sigma_i^2}$ and $w_0 = \frac{\mu_{i0}^2-\mu_{i1}^2}{\sigma_i^2}$ and plug back to equation (3), we can effectively get that

$$P(y = 1 \mid \mathbf{x}) = \frac{1}{1 + e^{w_0 + \sum_{i=1}^{n} w_i x_i}}$$

References

MITCHELL, T. M. (1997). *Machine Learning.* 1st ed. McGraw-Hill, Inc., New York, NY, USA.