Maschinelles Lernen 1 - Assignment 2

Technische Universität Berlin

Christoph Conrads (315565) Antje Relitz (327289) Benjamin Pietrowicz (332542)

Mitja Richter (324680)

WS 2013/2014

1 Gauging the Risk

$$
\begin{aligned}
f &= \text{a house will be flooded} & \neg f &= \text{a house won't be flooded} \\
x &= \text{house stands in a high risk area} & \neg x &= \text{house stands in a low risk area} \\
P(f) &= 0.0005 & P(\neg f) &= 0.9995 \\
P(x) &= 0.04 & P(\neg x) &= 0.96 \\
P(x|f) &= 0.8 & P(\neg x|f) &= 0.2 \\
\alpha_1 &= \text{buying insurance} & \alpha_2 &= \text{not buying insurance} \\
\lambda(\alpha_1|f) &= 1100€ & \lambda(\alpha_1|\neg f) &= 1100€ \\
\lambda(\alpha_2|f) &= 100000€ & \lambda(\alpha_2|\neg f) &= 0€
\end{aligned}
$$

a)

Bayes' formula is

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

We want to determine the probability that a house in a high risk area will be flooded, i. e., we want to calculate $P(f|x)$:

$$P(f|x) = \frac{P(x|f)P(f)}{P(x)} = \frac{0.8 \cdot 0.0005}{0.04} = 0.01$$

There is a 1% probability that our house will be flooded next year.

b)

We want to compute how much loss we can expect depending on whether we decide to buy insurance or not. Therefore we compute $R(\alpha_1|x)$ and $R(\alpha_2|x)$:

$$
\begin{aligned}
P(\neg f|x) &= 1 - P(f|x) = 0.99 \\
R(\alpha_1|x) &= \lambda(\alpha_1|f)P(f|x) + \lambda(\alpha_1|\neg f)P(\neg f|x) = 1100€ \cdot 0.01 + 1100€ \cdot 0.99 = 1100€ \\
R(\alpha_2|x) &= \lambda(\alpha_2|f)P(f|x) + \lambda(\alpha_2|\neg f)P(\neg f|x) = 100000€ \cdot 0.01 + 0€ \cdot 0.99 = 1000€
\end{aligned}
$$

Since $R(\alpha_2|x) < R(\alpha_1|x)$ it would be cheaper not to buy an insurance.

c)

We feel that the existing calculation matches this group's common sense.

2 Bounds on the Error

a)

Proof. We have

$$\min[P(\omega_1|x), P(\omega_2)|x)] \overset{!}{\leq} 2P(\omega_1|x)P(\omega_2|x)$$

WLOG let $P(\omega_1|x) \leq P(\omega_2|x)$ so that $0 \leq P(\omega_1|x) \leq 0.5$. Then

$$P(\omega_1|x) \leq 2P(\omega_1|x)P(\omega_2|x).$$

It holds that $P(\omega_2|x) = 1 - P(\omega_1|x)$ so

$$
\begin{aligned}
P(\omega_1|x) &\leq 2P(\omega_1|x)(1 - P(\omega_1|x)) \\
&\leq 2P(\omega_1|x) - 2P(\omega_1|x)^2 \Leftrightarrow \\
2P(\omega_1|x)^2 &\leq P(\omega_1|x)
\end{aligned}
$$

For $P(\omega_1|x) = 0$, the inequality holds. Otherwise we can divide by that term and get

$$P(\omega_1|x) \leq 0.5$$

which is fulfilled because of our assumptions. □

b)

Let $P(\omega_1|x) = P(\omega_2|x) = 0.5$. Then

$$
\begin{aligned}
P(\text{error}|x) = \min[P(\omega_1|x, P(\omega_2|x)] &\overset{!}{\leq} \alpha P(\omega_1|x)P(\omega_2|x) \\
0.5 &\leq 0.25\alpha \Leftrightarrow \\
\alpha &\geq 2
\end{aligned}
$$

which violates the assumption $\alpha < 2$.

3 Gaussian Densities

a)

We are looking for $P(\omega_1)$, $P(\omega_2)$ so that $P(\text{error}|x) = P(\omega_2|x)$. Let us assume that the latter equation is fulfilled then we know that $P(\omega_2|x) \leq P(\omega_1|x)$ because of the definition of $P(\text{error}|x)$. Using Bayes' rule, it holds that

$$\frac{P(x|\omega_2)P(\omega_2)}{\cancel{P(x)}} \leq \frac{P(x|\omega_1)P(\omega_1)}{\cancel{P(x)}}$$

We reformulate this to

$$\frac{P(x|\omega_2)}{P(x|\omega_1)} \leq \frac{P(\omega_1)}{P(\omega_2)} \tag{1}$$

and using the definitions we get

$$\frac{\exp(-(x+\mu)^2/(2\sigma^2))}{\exp(-(x-\mu)^2/(2\sigma^2))} \leq \frac{P(\omega_1)}{P(\omega_2)}$$

Applying the logarithm gives

$$\frac{1}{2\sigma^2}\left((x-\mu)^2 - (x+\mu)^2\right) = \frac{1}{2\sigma^2}(-4x\mu) \leq \ln\left(\frac{P(\omega_1)}{P(\omega_2)}\right)$$

Then we can conclude

$$\ln P(\omega_2) \leq \ln P(\omega_1) + \frac{2x\mu}{\sigma^2} \Leftrightarrow P(\omega_2) \leq e^{2x\mu/\sigma^2} P(\omega_1)$$

b)

We assume that $P(x|\omega_1) = ce^{-|x-\mu_1|/\sigma^2}, \mu_1 \in \mathbb{R}$ and $P(x|\omega_2) = ce^{-|x-\mu_2|/\sigma^2}, \mu_2 \in \mathbb{R}$ where $c \in \mathbb{R}$. To simplify the discussion we assume that $\mu_1 \leq \mu_2$. If we insert this into inequality 1, then

$$\frac{P(x|\omega_2)}{P(x|\omega_1)} = \frac{\exp(-|x-\mu_2|/\sigma^2)}{\exp(-|x-\mu_1|/\sigma^2)} = \exp\left(\frac{|x-\mu_1| - |x-\mu_2|}{\sigma^2}\right)$$

For different x we get

$$\ln\left(\frac{P(x|\omega_2)}{P(x|\omega_1)}\right) = \begin{cases} \frac{+\mu_1-\mu_2}{\sigma^2}, & \text{if } x \leq \mu_1 \\ \frac{-\mu_1-\mu_2+2x}{\sigma^2}, & \text{if } \mu_1 < x < \mu_2 \\ \frac{-\mu_1+\mu_2}{\sigma^2}, & \text{if } x \geq \mu_2 \end{cases}$$

So if x is outside of the open interval (μ_1, μ_2) then $P(\omega_1)/P(\omega_2)$ is not dependent on the exact value of x. If we assume additionally $\mu = \mu_2 = -\mu_1, \mu > 0$ for the Laplacian distribution then

$$P(\omega_2) \leq \begin{cases} e^{-2\mu/\sigma^2} P(\omega_1), & \text{if } x \leq -\mu \\ e^{2x/\sigma^2} P(\omega_1), & \text{if } -\mu < x < +\mu \\ e^{+2\mu/\sigma^2} P(\omega_1), & \text{if } x \geq +\mu \end{cases}$$