Machine Learning 1
Exercise Sheet 1

TU Berlin
WS 2017/18

Group: BSSBCH

Matthias Bigalke, 339547, maku@win.tu-berlin.de

Tolga Buz, 346836, buz_tolga@yahoo.de

Alejandro Hernandez, 395678, alejandrohernandezmunuera@gmail.com

Aitor Palacios Cuesta, 396276, aitor.palacioscuesta@campus.tu-berlin.de

Christof Schubert, 344450, christof.schubert@campus.tu-berlin.de

Daniel Steinhaus, 342563, dany.steinhaus@googlemail.com

Exercise 1: Estimating Bayes Error

(a) Show that the full error can be upper-bounded as follows:

$P(error) \leq \int \frac{2}{\frac{1}{P(\omega_1|x)}+\frac{1}{P(\omega_2|x)}}\ p(x)\ dx$

$\Leftrightarrow \int P(error|x)p(x)dx \leq \int \frac{2}{\frac{1}{P(\omega_1|x)}+\frac{1}{P(\omega_2|x)}}\ p(x)\ dx$

$\Leftrightarrow \int P(error|x)\ p(x)\ dx \leq \int \frac{2 \cdot P(\omega_1|x) \cdot P(\omega_2|x)}{P(\omega_2|x)+P(\omega_1|x)}\ p(x)\ dx$

Using Bayes formula $P(\omega_i|x) = \frac{p(x|\omega_i) \cdot P(\omega_i)}{p(x)}$ leads to:

$\Leftrightarrow \int P(error|x)p(x)dx \leq \int \frac{2 \cdot P(\omega_1|x) \cdot P(\omega_2|x) \cdot p(x)}{\frac{p(x|\omega_1) \cdot P(\omega_1)}{p(x)}+\frac{p(x|\omega_2) \cdot P(\omega_2)}{p(x)}}\ dx$

$\Leftrightarrow \int P(error|x)p(x)dx \leq \int \frac{2 \cdot P(\omega_1|x) \cdot P(\omega_2|x) \cdot p(x)^2}{p(x|\omega_1) \cdot P(\omega_1)+p(x|\omega_2) \cdot P(\omega_2)}\ dx$

With $p(x|\omega_1) \cdot P(\omega_1) + p(x|\omega_2) \cdot P(\omega_2) = p(x)$:

$\Leftrightarrow \int P(error|x)p(x)dx \leq \int 2 \cdot P(\omega_1|x) \cdot P(\omega_2|x) \cdot p(x)\ dx$

$\Leftrightarrow \int min[P(\omega_1|x), P(\omega_2|x)]p(x)dx \leq \int 2 \cdot P(\omega_1|x) \cdot P(\omega_2|x) \cdot p(x)\ dx$

If $f(x) \leq g(x)\ \forall\ x$, then $\int f(x)\ dx \leq \int g(x)\ dx\ \forall x$. Thus the integrals can be ignored:

$\Leftrightarrow min[P(\omega_1|x), P(\omega_2|x)]p(x) \leq 2 \cdot P(\omega_1|x) \cdot P(\omega_2|x) \cdot p(x)$

Now there are two cases to consider which can be handled the same way. First we are looking at case $P(\omega_1|x) \leq P(\omega_1|x)$:

$\Leftrightarrow P(\omega_1|x) \cdot p(x) \leq 2 \cdot P(\omega_1|x) \cdot P(\omega_2|x)p(x)$

$\Leftrightarrow 1 \leq 2 \cdot P(\omega_2|x)$

This is true, since $P(\omega_2|x) \geq P(\omega_1|x)$ and $P(\omega_1|x) + P(\omega_2|x) = 1$. Thus $P(\omega_2|x) \geq 0.5$ and $\Leftrightarrow 1 \leq 2 \cdot P(\omega_2|x)$ is valid.

The second case for $P(\omega_1|x) \geq P(\omega_1|x)$ can be solved analogously (and $1 \leq 2 \cdot P(\omega_1|x)$ is valid if $P(\omega_1|x) \geq P(\omega_2|x)$).

This shows that the full error can be upper-bound as given.

(b) Show that $P(error) \leq \frac{2P(\omega_1)P(\omega_2)}{\sqrt{1+4\mu^2 P(\omega_1)P(\omega_2)}}$

We are starting with the term of 1 a) which has been shown to be true:

$P(error) \leq \int \frac{2}{\frac{1}{P(\omega_1|x)}+\frac{1}{P(\omega_2|x)}}\ p(x)\ dx$

$\Leftrightarrow P(error) \leq \int \frac{2p(x)}{\frac{p(x)}{p(x|\omega_1)P(\omega_1)}+\frac{p(x)}{p(x|\omega_2)P(\omega_2)}}\ dx$

$\Leftrightarrow P(error) \leq \int \frac{2}{\frac{1}{p(x|\omega_1)P(\omega_1)}+\frac{1}{p(x|\omega_2)P(\omega_2)}}\ dx$

$\Leftrightarrow P(error) \leq \int \frac{2 \cdot p(x|\omega_1)P(\omega_1) \cdot p(x|\omega_2)P(\omega_2)}{p(x|\omega_1)P(\omega_1)+p(x|\omega_2)P(\omega_2)}\ dx$

$\Leftrightarrow P(error) \leq 2 \cdot P(\omega_1) \cdot P(\omega_2) \int \frac{1}{\pi^2 \cdot (1+(x-\mu)^2)(1+(x+\mu)^2) \cdot (\frac{P(\omega_1)}{\pi(1+(x-\mu)^2)}+\frac{P(\omega_2)}{\pi(1+(x+\mu)^2)})}\ dx$

$\Leftrightarrow P(error) \leq 2 \cdot P(\omega_1) \cdot P(\omega_2) \int \frac{1}{\pi(P(\omega_1) \cdot (1+(x+\mu)^2)+P(\omega_2) \cdot (1+(x-\mu)^2)}\ dx$

$\Leftrightarrow P(error) \leq \frac{2 \cdot P(\omega_1) \cdot P(\omega_2)}{\pi} \int \frac{1}{P(\omega_1)+P(\omega_1)(x^2+2\mu x+\mu^2)+P(\omega_2)+P(\omega_2)(x^2-2\mu x+\mu^2)}\ dx$

$\Leftrightarrow P(error) \leq \frac{2 \cdot P(\omega_1)P(\omega_2)}{\pi} \cdot \int \frac{1}{(P(\omega_1)+P(\omega_2))x^2+(2\mu(P(\omega_1)-P(\omega_2))x+\mu^2 P(\omega_1)+\mu^2 P(\omega_2)+P(\omega_1)+P(\omega_2)}$

Using the identity $\int \frac{1}{ax^2+bx+c}dx = \frac{2\pi}{\sqrt{4ac-b^2}}$ *:

$\Leftrightarrow P(error) \leq \frac{2 \cdot P(\omega_1)P(\omega_2)}{\pi} \cdot \frac{2\pi}{\sqrt{4 \cdot (P(\omega_1)+P(\omega_2)) \cdot (1+\mu^2)(P(\omega_1)+P(\omega_2))-4\mu^2(P(\omega_1)-P(\omega_2))^2}}$

$\Leftrightarrow P(error) \leq \frac{2 \cdot P(\omega_1)P(\omega_2)}{\cancel{\pi}} \cdot \frac{\cancel{2}\cancel{\pi}}{\cancel{2}\sqrt{(P(\omega_1)+P(\omega_2)) \cdot (1+\mu^2)(P(\omega_1)+P(\omega_2))-4\mu^2(P(\omega_1)-P(\omega_2))^2}}$

$\Leftrightarrow P(error) \leq \frac{2 \cdot P(\omega_1)P(\omega_2)}{\sqrt{(P(\omega_1)^2+2 \cdot P(\omega_1)P(\omega_2)+P(\omega_2)^2+4\mu^2 P(\omega_1)P(\omega_2)}}$

$\Leftrightarrow P(error) \leq \frac{2 \cdot P(\omega_1)P(\omega_2)}{\sqrt{(P(\omega_1)+P(\omega_2))^2+4\mu^2 P(\omega_1)P(\omega_2)}}$

With $(P(\omega_1)+P(\omega_2))^2 = 1$:

$\Leftrightarrow P(error) \leq \frac{2 \cdot P(\omega_1)P(\omega_2)}{\sqrt{1+4\mu^2 P(\omega_1)P(\omega_2)}} \qquad q.e.d.$

(c)

Exercise 2: Bayes Decision Boundaries

(a) The decision boundary will be:

$P(\omega_1|x) = P(\omega_2|x)$

$\Leftrightarrow \frac{p(x|\omega_1) \cdot P(\omega_1)}{\cancel{p}(x)} = \frac{p(x|\omega_2) \cdot P(\omega_2)}{\cancel{p}(x)}$

$\Leftrightarrow \frac{\cancel{1}}{\cancel{2}\sigma} \cdot exp(\frac{-(x-\mu)}{\sigma}) \cdot P(\omega_1) = \frac{\cancel{1}}{\cancel{2}\sigma} \cdot exp(\frac{-(x+\mu)}{\sigma}) \cdot P(\omega_2)$

$\Leftrightarrow \frac{|x-\mu|}{\sigma} + ln(P(\omega_1)) = \frac{|x+\mu|}{\sigma} + ln(P(\omega_2))$

*For using the integration identity formula, $b^2 \leq 4ac$ has to be ensured. This is given in our case, since $(1 + 4\mu^2 P(\omega_1)P(\omega_2)) > 0$

(b) That means: $P(\omega_1|x) > P(\omega_2|x) \; \forall \; x \in \mathbb{R}$

$\Leftrightarrow \frac{\cancel{1}}{\cancel{2}\sigma} \cdot exp(\frac{-(x-\mu)}{\sigma}) \cdot P(\omega_1) > \frac{\cancel{1}}{\cancel{2}\sigma} \cdot exp(\frac{-(x+\mu)}{\sigma}) \cdot P(\omega_2)$

$\Leftrightarrow ln(P(\omega_1)) - \frac{|x-\mu|}{\sigma} > ln(P(\omega_2)) - \frac{|x+\mu|}{\sigma}$

$\Leftrightarrow ln(P(\omega_1)) > ln(P(\omega_2)) - \frac{|x+\mu|}{\sigma} + \frac{|x-\mu|}{\sigma}$

Case: $\mu > 0 \Rightarrow \omega_1$ is selected when $P(\omega_1) > P(\omega_2)$

Case: $P(\omega_1) = 1 \Rightarrow \omega_1$ is always selected.

General case: ω_i is always selected when:
$ln(P\omega_1)) > ln(P(\omega_{"})) + \frac{2\mu}{\sigma}$
The right-hand side of the equality is the maximum of the function, achieved when $x = \mu$

(c) The derivation is equivalent:

Boundary:

$$ln(P(\omega_1)) - \frac{(x-\mu)^2}{2\sigma^2} = ln(P(\omega_2)) - \frac{(x+\mu)^2}{2\sigma^2}$$

Values:

$$ln(P(\omega_1)) - \frac{(x-\mu)^2}{2\sigma^2} > ln(P(\omega_2)) - \frac{(x+\mu)^2}{2\sigma^2}$$

$$\Leftrightarrow ln(P(\omega_1)) > ln(P(\omega_2)) - \frac{(x+\mu)^2}{2\sigma^2} + \frac{(x-\mu)^2}{2\sigma^2}$$

ω_1 is selected $\forall x \in \mathbb{R}$ when:

$\mu = 0$ and $P(\omega_1) > P(\omega_2)$

Otherwise, $ln(P(\omega_2)) - \frac{(x+\mu)^2}{2\sigma^2} + \frac{(x-\mu)^2}{2\sigma^2} \to \infty$ when $x \to -\infty$.

Exercise 3: Programming See Exercise1/sheet01.ipynb for the solution.