Machine Learning 1
Exercise Sheet 1

Wintersemester 2019/20

Gruppe:

Ben Aba, Mondher (TU Berlin, 380190)
Baga, Oleksandra (Beuth Hochshule für Technik, 849852)
Herrmannsdörfer, Daniel Viladrich (TU Berlin, 412543)
Janßen, Knut Arne (TU Berlin, 344086)
Kravets, Daria (Freie Universität Berlin, 4675700)
Tawfek, Noureldin (TU Berlin, 369213)

Abteilung Maschinelles Lernen
Institut für Softwaretechnik und theoretische Informatik
Fakultät IV, Technische Universität Berlin
Prof. Dr. Klaus-Robert Müller

22. Oktober 2019

1 Exercise 1: Estimating the Bayes Error

a)
The Bayes decision rule for the two classes classification problem results in the Bayes error

$$P(error) = \int P(error|x)p(x)dx$$

We have to show that the full error can be upper-bounded as follows:

$$P(error) \le \int \frac{2}{\frac{1}{P(w_1|x)} + \frac{1}{P(w_2|x)}} p(x)dx$$

we know from the question that

$$\int P(error|x)p(x)dx \le \int \frac{2}{\frac{1}{P(w_1|x)} + \frac{1}{P(w_2|x)}} p(x)dx$$

We can show only the equality:

$$P(error|x) \le \frac{2}{\frac{1}{P(w_1|x)} + \frac{1}{P(w_2|x)}}$$

As we know from the task and lecture:

$$P(error|x) = min[P(w_1|x), P(w_2|x)]$$

So we can show to the equality of the following:

$$\begin{aligned}
min[P(w_1|x), P(w_2|x)] &\le \frac{2}{\frac{1}{P(w_1|x)} + \frac{1}{P(w_2|x)}} \\
min[P(w_1|x), P(w_2|x)] \times \left(\frac{1}{P(w_1|x)} + \frac{1}{P(w_2|x)}\right) &\le 2 \\
min\left[\cancel{\frac{P(w_1|x)}{P(w_1|x)}} + \frac{P(w_1|x)}{P(w_2|x)}, \cancel{\frac{P(w_2|x)}{P(w_2|x)}} + \frac{P(w_2|x)}{P(w_1|x)}\right] &\le 2 \\
min\left[1 + \frac{P(w_1|x)}{P(w_2|x)}, 1 + \frac{P(w_2|x)}{P(w_1|x)}\right] &\le 2 \\
min\left[\frac{P(w_1|x)}{P(w_2|x)}, \frac{P(w_2|x)}{P(w_1|x)}\right] &\le 1
\end{aligned}$$

we know The Bayes decision rule for the two classes:

$$= \begin{cases} w_1, & \text{if } P(w_1|x) \ge P(w_2|x) \\ w_2, & \text{if } P(w_2|x) \ge P(w_1|x) \end{cases}$$

So we can take 3 variants for our problem:
1) if we have decided w_1,

$$min\left[\frac{P(w_1|x)}{P(w_2|x)}, \frac{P(w_2|x)}{P(w_1|x)}\right] = \frac{P(w_2|x)}{P(w_1|x)} < 1$$

2) if we have decided w_2,

$$min\left[\frac{P(w_1|x)}{P(w_2|x)}, \frac{P(w_2|x)}{P(w_1|x)}\right] = \frac{P(w_1|x)}{P(w_2|x)} < 1$$

3) and if $P(w_1|x) = P(w_2|x)$,

$$min\left[\frac{P(w_1|x)}{P(w_2|x)}, \frac{P(w_2|x)}{P(w_1|x)}\right] = 1$$

Which was to be proven.

b)
We start with the inequality derived in section a)

$$P(error) \leq \int \frac{2}{\frac{1}{p(w_1|x)} + \frac{1}{p(w_2|x)}} p(x)dx$$

The rest of the exercise is manipulating this expression until it has the form given in the exercise: We apply Bayes' rule to the conditional Probabilities $p(w_i|x) = \frac{p(x|w_i)*P(w_i)}{p(x)}$

$$P(error) \leq 2\int \frac{1}{\frac{p(x)}{p(x|w_1)P(w_1)} + \frac{p(x)}{p(x|w_2)P(w_2}} p(x)dx$$
$$P(error) \leq 2\int \frac{1}{\frac{1}{p(x|w_1)P(w_1)} + \frac{1}{p(x|w_2)P(w_2}} dx$$
$$P(error) \leq 2\int \frac{1}{\frac{p(x|w_1)P(w_1)+p(x|w_2)P(w_2}{p(x|w_1)P(w_1)p(x|w_2)P(w_2}} dx$$
$$P(error) \leq 2\int \frac{p(x|w_1)P(w_1)p(x|w_2)P(w_2)}{p(x|w_1)P(w_1) + p(x|w_2)P(w_2)} dx$$
$$P(error) \leq 2\int \frac{P(w_1)P(w_2)}{\frac{p(x|w_2)P(w_2)}{p(x|w_1)p(x|w_2)} + \frac{p(x|w_1)P(w_1)}{p(x|w_1)p(x|w_2)}} dx$$
$$P(error) \leq 2\int \frac{P(w_1)P(w_2)}{\frac{P(w_2)}{p(x|w_1)} + \frac{P(w_1)}{p(x|w_2)}} dx$$
$$P(error) \leq \frac{2P(w_1)P(w_2)}{\pi} \int \frac{1}{P(w_2)(1+(x-\mu)^2) + P(w_1)(1+(x+\mu)^2)} dx$$
$$P(error) \leq \frac{2P(w_1)P(w_2)}{\pi} \int \frac{1}{(P(w_1)+P(w_2))x^2 + 2(P(w_1)-P(w_2))\mu x + (P(w_1)+P(w_2))(1+\mu^2)} dx$$

We use the identity $P(w_1) + P(w_2) = 1$ since the two events are complementary

$$P(error) \leq \frac{2P(w_1)P(w_2)}{\pi} \int \frac{1}{x^2 + 2(P(w_1)-P(w_2))\mu x + (1+\mu^2)} dx$$
$$P(error) \leq \frac{2P(w_1)P(w_2)}{\pi} \frac{2\pi}{\sqrt{4(1+\mu^2) - 4(P(w_1)-P(w_2))^2}}$$
$$P(error) \leq \frac{2P(w_1)P(w_2)}{\sqrt{(1+\mu^2) - \mu^2(P(w_1)-P(w_2))^2}}$$

Note the identity:$(P(w_1) - P(w_2))^2 = (P(w_1) + P(w_2))^2 - 4P(w_1)P(w_2) = 1 - 4P(w_1)P(w_2)$

$$P(error) \leq \frac{2P(w_1)P(w_2)}{\sqrt{(1+\mu^2) - \mu^2(1-4P(w_1)P(w_2))}}$$
$$P(error) \leq \frac{2P(w_1)P(w_2)}{\sqrt{1+4\mu^2 P(w_1)P(w_2)}}$$

Which was to be proven.

c)
Because it is known that the Bayes decision rule for the two classes classification problem results in the Bayes error $P(error) = \int P(error|x)p(x)dx$ that means the same as sum of all probabilities of $P(error, x)$. Without the upper-bounds that are both tight and analytically integrable we would use the sum of all $P(error, x)$.

2 Exercise 2: Bayes Decision Boundaries

a)
The Bayes error will reach the maximum when $P(w_1|x) = P(w_2|x)$, because the both classes is equally the same to be decided.

$$P(w_1|x) = \frac{P(x|w_1)P(w_1)}{P(x)}$$

$$P(w_2|x) = \frac{P(x|w_2)P(w_2)}{P(x)}$$

$$\frac{P(x|w_1)P(w_1)}{P(x)} = \frac{P(x|w_2)P(w_2)}{P(x)}$$

$$\frac{P(x|w_2)}{P(x|w_1)} = \frac{\exp\left(\frac{-|x+\mu|}{\sigma}\right)}{\exp\left(\frac{-|x-\mu|}{\sigma}\right)}$$

$$\Leftrightarrow \exp\left(-\frac{|\ x-\mu\ |}{\sigma}\right) P(w_1) = \exp\left(-\frac{|\ x+\mu\ |}{\sigma}\right) P(w_2)$$

To solve an exponential equation, we take the log of both sides, and solve for the variable.

$$\begin{aligned}
\Leftrightarrow \qquad & \ln(\exp\left(-\frac{|\ x-\mu\ |}{\sigma}\right) P(w_1)) = \ln(\exp\left(-\frac{|\ x+\mu\ |}{\sigma}\right) P(w_2)) \\
\Leftrightarrow \qquad & -\frac{|\ x-\mu\ |}{\sigma} + \ln(P(w_1)) = -\frac{|\ x+\mu\ |}{\sigma} + \ln(P(w_2)) \\
\Leftrightarrow \qquad & \ln(P(w_1)) - \ln(P(w_2)) = \frac{|\ x-\mu\ |}{\sigma} - \frac{|\ x+\mu\ |}{\sigma} \\
\Leftrightarrow \qquad & \ln\left(\frac{P(w_1)}{P(w_2)}\right) = \frac{|\ x-\mu\ | - |\ x+\mu\ |}{\sigma} \\
\Leftrightarrow \qquad & \sigma \ln\left(\frac{P(w_1)}{P(w_2)}\right) =|\ x-\mu\ | - |\ x+\mu\ |
\end{aligned}$$

We can take 4 cases to determine the boundaries:

1. $x < \mu$ and $-x < \mu$

$$\begin{aligned}
\sigma \ln\left(\frac{P(w_1)}{P(w_2)}\right) &= -(x-\mu) - (x+\mu) \\
&= -2x \\
-\frac{\sigma}{2} \ln\left(\frac{P(w_1)}{P(w_2)}\right) &= x
\end{aligned}$$

2. $x > \mu$ and $-x > \mu$

$$\begin{aligned}
\sigma \ln\left(\frac{P(w_1)}{P(w_2)}\right) &= x - \mu + x + \mu \\
&= 2x \\
\frac{\sigma}{2} \ln\left(\frac{P(w_1)}{P(w_2)}\right) &= x
\end{aligned}$$

3. $x \leq \mu$ and $-x \geq \mu$

$$\begin{aligned}
\sigma \ln\left(\frac{P(w_1)}{P(w_2)}\right) &= -(x-\mu) + x + \mu \\
&= 2\mu
\end{aligned}$$

4. $x \geq \mu$ and $-x \leq \mu$

$$\begin{aligned}
\sigma \ln\left(\frac{P(w_1)}{P(w_2)}\right) &= (x-\mu) - (x+\mu) \\
&= -2\mu
\end{aligned}$$

From the equations (3) and (4) we see, that the boundary must lie between the mean μ of both classes.

b

Nour's Solutions