Niklas Gruhn (389343), Haron Shaker (415349), Malek Alhelwany (391558)

Machine Learning - Exercise 1

Question 2

(a)

We want to calculate

$$p(apple) = p(apple|r)p(r) + p(apple|g)p(g) + p(apple|b)p(b)$$

The individual terms are calculated as follows

$$\begin{aligned} p(apple|r) &= \frac{3}{3+4+3} = \frac{3}{10} \\ p(apple|g) &= \frac{3}{3+3+4} = \frac{3}{10} \\ p(apple|b) &= \frac{1}{1+1+0} = \frac{1}{2} \end{aligned}$$

so we obtain

$$p(apple) = \frac{3}{10} \cdot 0.2 + \frac{3}{10} \cdot 0.6 + \frac{1}{2} \cdot 0.2 = 0.34$$

(b)

Here we want to calculate

$$p(g|orange) = \frac{p(orange|g)p(g)}{p(orange)}$$

Again we first need to figure out a few other terms

$$\begin{aligned} p(orange|r) &= \frac{4}{3+4+3} = \frac{2}{5} \\ p(orange|g) &= \frac{3}{3+3+4} = \frac{3}{10} \\ p(orange|b) &= \frac{1}{1+1+0} = \frac{1}{2} \end{aligned}$$

with that we can calculate

$$p(orange) = p(orange|r)p(r) + p(orange|g)p(g) + p(orange|b)p(b) = 0.08 + 0.18 + 0.1 = 0.36$$

Putting it all together we get

$$p(g|orange) = \frac{0.18}{0.36} = 0.5$$

Question 3

(a)

For any class $C_j \in C$ the expected loss is given by

$$\mathbb{E}[L] = \sum_k L_{k,j} p(C_k|x)$$

Using the definition of $L_{k,j}$ and the fact that $\sum_k p(C_k|x) = 1$ we can rewrite this equation as

$$\begin{aligned}\mathbb{E}[L] &= \sum_k L_{k,j} p(C_k|x) \\ &= 0 \cdot p(C_j|x) + l_s \sum_{k \neq j} p(C_k|x) \\ &= l_s(1 - p(C_j|x))\end{aligned}$$

Now suppose C_j is the class with minimum expected loss for the sample x. Then we have

$$\begin{aligned} & \forall k : l_s(1 - p(C_j|x)) \leq l_s(1 - p(C_k|x) \\ \iff \quad & \forall k : -p(C_j|x)) \leq -p(C_k|x) \\ \iff \quad & \forall k : p(C_j|x)) \geq p(C_k|x)\end{aligned}$$

The expected loss must also be smaller than the loss incurred from rejection l_r. So we also have

$$\begin{aligned} & l_s(1 - p(C_j|x)) \leq l_r \\ \iff \quad & 1 - p(C_j|x) \leq \frac{l_r}{l_s} \\ \iff \quad & p(C_j|x) \geq 1 - \frac{l_r}{l_s}\end{aligned}$$

Putting both results together we get

$$\forall k : p(C_j|x)) \geq p(C_k|x) \wedge p(C_j|x) \geq 1 - \frac{l_r}{l_s}$$

Thus we showed that C_j is the class with minimum expected loss iff the condition above is satisfied.

(b)

When $l_r = 0$ then rejecting is "loss-free" and the expected loss is minimized by just rejecting all actions unless the class can be determined with perfect certainty. More formally, the condition

$$p(C_j|x) \geq 1 - \frac{l_r}{l_s}$$

becomes

$$p(C_j|x) \geq 1$$

which is only true when $p(C_j|x) = 1$. So we get

$$\text{classify}(x) :\rightarrow \begin{cases} C_j & \text{if} \quad p(C_j|x) = 1 \\ C_{rej} & \text{otherwise} \end{cases}$$

(c)

When $l_r > l_s$ then rejecting entails greater loss than any substitution error. Hence samples are never rejected. Formally, we can see that the right-hand side of this inequality

$$p(C_j|x) \geq 1 - \frac{l_r}{l_s}$$

is negative so the condition is always true. The second condition

$$\forall k : p(C_j|x)) \geq p(C_k|x)$$

is also always true for some C_j. Thus we get

$$\text{classify}(x) :\to C_j \quad \text{with} \quad \forall k : p(C_j|x)) \geq p(C_k|x)$$

Question 4

We have to find θ for which

$$p(X|\theta) = \prod_{n=0}^{N} p(x_n|\theta)$$

is maximal. Or equivalently, minimize

$$-ln(p(X|\theta)) = -\sum_{n=0}^{N} ln(p(x_n|\theta))$$

For that we compute the derivative.

$$\begin{aligned}
&-\frac{\partial}{\partial\theta}\sum_{n=0}^{N} ln(p(x_n|\theta)) \\
=&-\frac{\partial}{\partial\theta}\sum_{n=0}^{N}(ln(\theta^2 x_n exp(-\theta x_n)g(x_n))
\end{aligned}$$

Since all data points are defined to be greater zero, we can remove the term $g(x_n)$.

$$\begin{aligned}
&-\frac{\partial}{\partial\theta}\sum_{n=0}^{N}(ln(\theta^2 x_n exp(-\theta x_n)) \\
=&-(\sum_{n=0}^{N}(\frac{\partial}{\partial\theta}ln(\theta^2) + \frac{\partial}{\partial\theta}ln(x_n) - \frac{\partial}{\partial\theta}\theta x_n)) \\
=&-(\sum_{n=0}^{N}(\frac{2}{\theta} - x_n)) \\
=&-\frac{2N}{\theta} + \sum_{n=0}^{N} x_n
\end{aligned}$$

By equating to zero and solving for θ, we obtain the he maximum likelihood estimate

$$\begin{aligned}
&-\frac{2N}{\theta} + \sum_{n=0}^{N} x_n = 0 \\
\iff \quad &\theta = \frac{2N}{\sum_{n=0}^{N} x_n}
\end{aligned}$$