Exercise 2

Machine Learning in Graphics & Vision

Driton Goxhufi 4233242 driton.goxhufi@student.uni-tuebingen.de
Damir Ravlija 5503184 damir.ravlija@student.uni-tuebingen.de

1 Task 1

(a) Classification accuracy of the initialized model on the test dataset is 0.5

(b) Loss of the initialized model is 0.7149616252170096.

(c) In the first step of derivation we use the chain rule and the fact that $f'_{\boldsymbol{w}}(x) = f_{\boldsymbol{w}}(x)(1 - f_{\boldsymbol{w}}(x)$.

$$\begin{aligned}
\frac{\partial L(\boldsymbol{x}, t, \boldsymbol{w})}{\partial \boldsymbol{w}} &\overset{(1)}{=} \frac{1}{N}\sum_{n=1}^{N}\left[-t_n \frac{1}{f_{\boldsymbol{w}}(\boldsymbol{x}_n)} f_{\boldsymbol{w}}(\boldsymbol{x}_n)(1 - f_{\boldsymbol{w}}(\boldsymbol{x}_n))\boldsymbol{x}_n + (1 - t_n)\frac{1}{1 - f_{\boldsymbol{w}}(\boldsymbol{x_n})} f_{\boldsymbol{w}}(\boldsymbol{x}_n)(1 - f_{\boldsymbol{w}}(\boldsymbol{x}_n))\boldsymbol{x}_n\right] \\
&= \frac{1}{N}\sum_{n=1}^{N}\left[-t_n(1 - f_{\boldsymbol{w}}(\boldsymbol{x}_n))\boldsymbol{x}_n + (1 - t_n) f_{\boldsymbol{w}}(\boldsymbol{x}_n)\boldsymbol{x}_n\right] \\
&= \frac{1}{N}\sum_{n=1}^{N}\left[(-t_n + t_n f_{\boldsymbol{w}}(\boldsymbol{x}_n) + f_{\boldsymbol{w}}(\boldsymbol{x}_n) - t_n f_{\boldsymbol{w}}(\boldsymbol{x}_n))\boldsymbol{x}_n\right] \\
&= \frac{1}{N}\sum_{n=1}^{N}\left[f_{\boldsymbol{w}}(\boldsymbol{x}_n) - t_n\right]\boldsymbol{x}_n
\end{aligned}$$

(1)

$$\begin{aligned}
\frac{\partial f_{\boldsymbol{w}}(x)}{\partial x} &= \frac{\partial}{\partial x}\left(\frac{1}{1 + e^{-\boldsymbol{w}^T\boldsymbol{x}}}\right) = \frac{e^{-\boldsymbol{w}^T\boldsymbol{x}}}{\left(1 + e^{-\boldsymbol{w}^T\boldsymbol{x}}\right)^2} = \frac{1 + e^{-\boldsymbol{w}^T\boldsymbol{x}} - 1}{\left(1 + e^{-\boldsymbol{w}^T\boldsymbol{x}}\right)^2} = \frac{1}{\left(1 + e^{-\boldsymbol{w}^T\boldsymbol{x}}\right)} - \frac{1}{\left(1 + e^{-\boldsymbol{w}^T\boldsymbol{x}}\right)^2} \\
&= \frac{1}{\left(1 + e^{-\boldsymbol{w}^T\boldsymbol{x}}\right)}\left(1 - \frac{1}{\left(1 + e^{-\boldsymbol{w}^T\boldsymbol{x}}\right)}\right) = f_{\boldsymbol{w}}(\boldsymbol{w}^T\boldsymbol{x})(1 - f_{\boldsymbol{w}}(\boldsymbol{w}^T\boldsymbol{x}))
\end{aligned}$$

After $1,000$ iterations the loss and accuracy of the model are:

loss = 0.3868595564299156

accuracy = 0.83