Assignment 1: Machine learning basics

Jan Scheffczyk - 3242317
Sarah Khan - 3279206
Mahmoud Hashem - 3201329
Mohamed Saleh - 3201337

October 30, 2019

1 Linear Classifier with Logistic Function

Let

$$\begin{aligned}
W &\quad \text{weights} \in \mathbb{R}^D \\
X &\quad \text{inputs} \in \mathbb{R}^{N\times D} \\
b &\quad \text{bias} \in \mathbb{R} \\
y &\quad \text{labels} \in \mathbb{B}^N \\
h\left(x_i, W, b\right) = &\quad \sigma\left(W x_i + b\right) \\
\sigma\left(a\right) = &\quad \frac{1}{1+\exp\left(-a\right)} \\
k = &\quad W x_i + b
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial w} L\left(w, b\right) &= \frac{\partial}{\partial w} \sum_{i=1}^{N} \left(y_i - h\left(x_i, W, b\right)\right)^2 \\
&= \sum_{i=1}^{N} 2\left(y_i - \sigma\left(k\right)\right)\left(-\right)\frac{\partial}{\partial w}\sigma\left(k\right) \\
&= \sum_{i=1}^{N} 2\left(y_i - \sigma\left(k\right)\right)\left(-\right)\left(\sigma\left(k\right)\right)\left(1-\sigma\left(k\right)\right)\frac{\partial}{\partial w}k \qquad \text{using}(1) \\
&= \sum_{i=1}^{N} 2\left(y_i - \sigma\left(k\right)\right)\left(-\right)\left(\sigma\left(k\right)\right)\left(1-\sigma\left(k\right)\right)x_i
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial}{\partial b} L\left(w, b\right) \quad &= \frac{\partial}{\partial b} \sum_{i=1}^{N} \left(y_i - h\left(x_i, W, b\right)\right)^2 \\
&= \sum_{i=1}^{N} 2\left(y_i - \sigma\left(k\right)\right)\left(-\right) \frac{\partial}{\partial b} \sigma\left(k\right) \\
&= \sum_{i=1}^{N} 2\left(y_i - \sigma\left(k\right)\right)\left(-\right)\left(\sigma\left(k\right)\right)\left(1 - \sigma\left(k\right)\right) \frac{\partial}{\partial b} k \qquad \text{using}(1) \\
&= \sum_{i=1}^{N} 2\left(y_i - \sigma\left(k\right)\right)\left(-\right)\left(\sigma\left(k\right)\right)\left(1 - \sigma\left(k\right)\right)\left(1\right)
\end{aligned}
$$

$$
\begin{aligned}
\frac{d}{dx}\sigma(x) &= \frac{d}{dx}\left[\frac{1}{1+e^{-x}}\right] \\
&= \frac{d}{dx}\left(1+\mathrm{e}^{-x}\right)^{-1} \\
&= -(1+e^{-x})^{-2}(-e^{-x}) \\
&= \frac{e^{-x}}{\left(1+e^{-x}\right)^2} \\
&= \frac{1}{1+e^{-x}} \cdot \frac{e^{-x}}{1+e^{-x}} \\
&= \frac{1}{1+e^{-x}} \cdot \frac{(1+e^{-x})-1}{1+e^{-x}} \\
&= \frac{1}{1+e^{-x}} \cdot \left(\frac{1+e^{-x}}{1+e^{-x}} - \frac{1}{1+e^{-x}}\right) \\
&= \frac{1}{1+e^{-x}} \cdot \left(1 - \frac{1}{1+e^{-x}}\right) \\
&= \sigma(x) \cdot (1 - \sigma(x))
\end{aligned} \tag{1}
$$