

Übungsaufgaben

51) Es wird ein künstliches Neuronales Netz betrachtet, das nur aus einem Input und einem Output Layer besteht (d.h. ein künstliches neuronales Netz mit $L = 1$ Schichten). Mit Hilfe der Inputvariablen

- Alter (x_1),
- Einkommen (x_2) und
- Anzahl der bisherigen Käufe (x_3)

sagt es für die Kunden eines Unternehmens vorher, ob Werbung zum Kauf eines bestimmten Produkts führt. Für die Outputvariable gilt dabei

$$t = \begin{cases} 0 & \text{falls das Produkt nicht gekauft wird} \\ 1 & \text{falls das Produkt gekauft wird} \end{cases}.$$

Der Bias und die Gewichte für die Inputvariablen wurden bereits mit Hilfe von Kundendaten geschätzt und sind gegeben durch

$$b = 0, \quad w_1 = -0,1, \quad w_2 = 0,6 \quad \text{und} \quad w_3 = 0,7.$$

Als Aktivierungsfunktion für die Ausgabeschicht wird die Logistische Funktion

$$\sigma(x) = \frac{1}{1 + \exp(-x)}$$

verwendet.

Übungsaufgaben

Bestimmen Sie für die folgenden Kunden die Wahrscheinlichkeit (gerundet auf vier Nachkommastellen), dass sie das Produkt kaufen:

	Kunde 1	Kunde 2	Kunde 3
Alter x_1	20	30	40
Einkommen x_2	6	5	1
Anzahl bisheriger Käufe x_3	1	0	3

Interpretieren Sie ferner die erhaltenen Ergebnisse.

Übungsaufgaben

- 56) Betrachtet wird ein zweischichtiges Feedforward-Netz (d.h. $L = 2$) mit $d_0 = d_1 = 2$, $d_2 = 1$, den Aktivierungsfunktionen

$$\sigma^{(1)}(z) = \text{ReLU}(z) = \max\{z, 0\} \quad \text{und} \quad \sigma^{(2)}(z) = z$$

sowie der Verlustfunktion

$$C(\boldsymbol{\theta}) = L(\mathbf{y}_{\boldsymbol{\theta}}(\mathbf{x}), t) = (\mathbf{y}_{\boldsymbol{\theta}}(\mathbf{x}) - t)^2.$$

Verwenden Sie die Definitionen

$$\mathbf{z}^{(l)} := \mathbf{W}^{(l)} \mathbf{a}^{(l-1)} + \mathbf{b}^{(l)}, \quad \mathbf{a}^{(l)} := \sigma^{(l)}(\mathbf{z}^{(l)}), \quad \mathbf{a}^{(0)} := \mathbf{x} \quad \text{und} \quad \mathbf{y}_{\boldsymbol{\theta}}(\mathbf{x}) := \mathbf{a}^{(2)}$$

sowie die Backpropagation-Formeln

$$\frac{\partial C(\boldsymbol{\theta})}{\partial \mathbf{W}^{(l)}} = \boldsymbol{\delta}^{(l)} \mathbf{a}^{(l-1)T} \in \mathbb{R}^{d_l \times d_{l-1}} \quad \text{und} \quad \frac{\partial C(\boldsymbol{\theta})}{\partial \mathbf{b}^{(l)}} = \boldsymbol{\delta}^{(l)} \in \mathbb{R}^{d_l} \quad \text{für } l = 1, 2$$

mit

$$\boldsymbol{\delta}^{(1)} = \sigma^{(1)'}(\mathbf{z}^{(1)}) \odot \mathbf{W}^{(2)T} \boldsymbol{\delta}^{(2)} \quad \text{und} \quad \boldsymbol{\delta}^{(2)} := \sigma^{(2)'}(\mathbf{z}^{(2)}) \odot \frac{\partial C(\boldsymbol{\theta})}{\partial \mathbf{y}_{\boldsymbol{\theta}}(\mathbf{x})}.$$

Gegeben seien

$$\mathbf{W}^{(1)} = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}, \quad \mathbf{b}^{(1)} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad \mathbf{W}^{(2)} = (1 \quad -2), \quad \mathbf{b}^{(2)} = \frac{1}{2}$$

und das Input-Output-Paar $\mathbf{x} = (1, 2)^T$ und $t = 3$.

Übungsaufgaben

- a) Feedforward-Schritt: Berechnen Sie $\mathbf{z}^{(1)}$, $\mathbf{a}^{(1)}$, $\mathbf{z}^{(2)}$, $\mathbf{y}_{\boldsymbol{\theta}}(\mathbf{x})$ und $C(\boldsymbol{\theta})$.
b) Backpropagation-Schritt: Berechnen Sie $\delta^{(2)}$ und $\delta^{(1)}$ sowie die Teilgradienten

$$\frac{\partial C(\boldsymbol{\theta})}{\partial \mathbf{W}^{(2)}}, \frac{\partial C(\boldsymbol{\theta})}{\partial \mathbf{b}^{(2)}}, \frac{\partial C(\boldsymbol{\theta})}{\partial \mathbf{W}^{(1)}} \quad \text{und} \quad \frac{\partial C(\boldsymbol{\theta})}{\partial \mathbf{b}^{(1)}}.$$

- c) Gradientenabstiegsverfahren: Berechnen Sie den ersten Schritt des Gradientenabstiegsverfahrens mit der Lernrate $\eta = 0,01$. D.h. berechnen Sie

$$\mathbf{W}_{neu}^{(l)} = \mathbf{W}^{(l)} - \eta \frac{\partial C(\boldsymbol{\theta})}{\partial \mathbf{W}^{(l)}}$$
$$\mathbf{b}_{neu}^{(l)} = \mathbf{b}^{(l)} - \eta \frac{\partial C(\boldsymbol{\theta})}{\partial \mathbf{b}^{(l)}}$$

für $l = 1, 2$ und $\eta = 0,01$. Bestimmen Sie anschließend den neuen Output $\mathbf{y}_{\boldsymbol{\theta}}(\mathbf{x})_{neu}$ und den neuen Verlust $C(\boldsymbol{\theta}_{neu})$.