

Mathematikaufgabe 95

[Home](#) | [Startseite](#) | [Impressum](#) | [Kontakt](#) | [Gästebuch](#)

Aufgabe: Übertragen Sie 8 neuronale Informationen durch einen Flaschenhals von 3 verdeckten Neuronen, ohne daß dabei die Information verändert wird.

Lösung: Der 8-Bit-Encoder ist ein Paradebeispiel für nichtüberwachtes Lernen. Wir lösen die Aufgabenstellung mit Hilfe eines dreilagigen Feed-Forward-Netzwerks (Abb. 1), wobei die Ausgangsaktivitäten durch die 3 Gleichungen

$$x_i^{(3)} = g\left(\sum_j w_{ij}^{(2)} x_j^{(2)}\right) = y_i$$
$$x_i^{(2)} = g\left(\sum_j w_{ij}^{(1)} x_j^{(1)}\right)$$
$$x_i^{(1)} \equiv x_i = \text{Eingangsknoten}$$

gegeben sind, mit den Gewichten $w_{ij}^{(k)}$. Unser Trainingsdatensatz besteht aus 8 Mustervektoren mit je 8 Eingangsknoten:¹

$$\mathbf{x}^{(p)} = (x_1^{(p)}, \dots, x_8^{(p)})$$

Für jedes Muster p und jeden Ausgangsknoten i ist die richtige Antwort $\hat{y}_i^{(p)}$ bekannt und kann daher mit der Antwort $y_i^{(p)}$ des Netzes verglichen werden. Das geschieht anhand der Fehlerfunktion

$$E = \frac{1}{2} \sum_{p=1}^8 \sum_{i=1}^8 (y_i^{(p)} - \hat{y}_i^{(p)})^2,$$

wobei die zweite Summation über die Zahl der Ausgangsknoten vorgenommen wird.

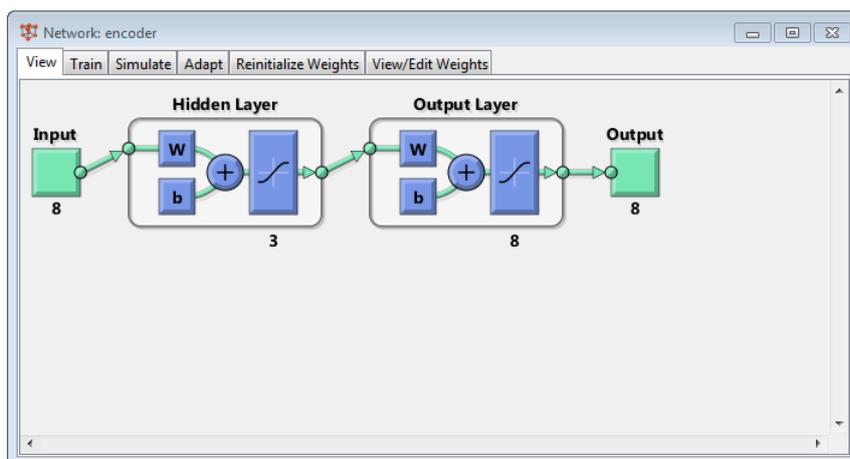


Abbildung 1. Schematische grafische Darstellung des verwendeten Netzwerks

¹ Den Index k lassen wir im folgenden weg, weil wir die Variable x immer nur für den Eingangslayer verwenden.

Mathematikaufgabe 95

Trainiert werden also konkret die folgenden 8 Mustervektoren, denen die Binärwerte

$$x_i^{(p)} = \delta_{ip} = \begin{cases} 1 & \text{falls } i = p \\ 0 & \text{sonst} \end{cases}$$

zugeordnet sind:

$$\begin{aligned} \mathbf{x}^{(1)} &= (x_1^{(1)}, x_2^{(1)}, x_3^{(1)}, x_4^{(1)}, x_5^{(1)}, x_6^{(1)}, x_7^{(1)}, x_8^{(1)}) = (1, 0, 0, 0, 0, 0, 0, 0) \\ \mathbf{x}^{(2)} &= (x_1^{(2)}, x_2^{(2)}, x_3^{(2)}, x_4^{(2)}, x_5^{(2)}, x_6^{(2)}, x_7^{(2)}, x_8^{(2)}) = (0, 1, 0, 0, 0, 0, 0, 0) \\ \mathbf{x}^{(3)} &= (x_1^{(3)}, x_2^{(3)}, x_3^{(3)}, x_4^{(3)}, x_5^{(3)}, x_6^{(3)}, x_7^{(3)}, x_8^{(3)}) = (0, 0, 1, 0, 0, 0, 0, 0) \\ \mathbf{x}^{(4)} &= (x_1^{(4)}, x_2^{(4)}, x_3^{(4)}, x_4^{(4)}, x_5^{(4)}, x_6^{(4)}, x_7^{(4)}, x_8^{(4)}) = (0, 0, 0, 1, 0, 0, 0, 0) \\ \mathbf{x}^{(5)} &= (x_1^{(5)}, x_2^{(5)}, x_3^{(5)}, x_4^{(5)}, x_5^{(5)}, x_6^{(5)}, x_7^{(5)}, x_8^{(5)}) = (0, 0, 0, 0, 1, 0, 0, 0) \\ \mathbf{x}^{(6)} &= (x_1^{(6)}, x_2^{(6)}, x_3^{(6)}, x_4^{(6)}, x_5^{(6)}, x_6^{(6)}, x_7^{(6)}, x_8^{(6)}) = (0, 0, 0, 0, 0, 1, 0, 0) \\ \mathbf{x}^{(7)} &= (x_1^{(7)}, x_2^{(7)}, x_3^{(7)}, x_4^{(7)}, x_5^{(7)}, x_6^{(7)}, x_7^{(7)}, x_8^{(7)}) = (0, 0, 0, 0, 0, 0, 1, 0) \\ \mathbf{x}^{(8)} &= (x_1^{(8)}, x_2^{(8)}, x_3^{(8)}, x_4^{(8)}, x_5^{(8)}, x_6^{(8)}, x_7^{(8)}, x_8^{(8)}) = (0, 0, 0, 0, 0, 0, 0, 1) \end{aligned}$$

Die zu erwartenden Netzantworten, für die analog

$$\hat{y}_i^{(p)} = \delta_{ip} = \begin{cases} 1 & \text{falls } i = p \\ 0 & \text{sonst} \end{cases}$$

gilt, lauten dann

$$\begin{aligned} \hat{\mathbf{y}}^{(1)} &= (\hat{y}_1^{(1)}, \hat{y}_2^{(1)}, \hat{y}_3^{(1)}, \hat{y}_4^{(1)}, \hat{y}_5^{(1)}, \hat{y}_6^{(1)}, \hat{y}_7^{(1)}, \hat{y}_8^{(1)}) = (1, 0, 0, 0, 0, 0, 0, 0) \\ \hat{\mathbf{y}}^{(2)} &= (\hat{y}_1^{(2)}, \hat{y}_2^{(2)}, \hat{y}_3^{(2)}, \hat{y}_4^{(2)}, \hat{y}_5^{(2)}, \hat{y}_6^{(2)}, \hat{y}_7^{(2)}, \hat{y}_8^{(2)}) = (0, 1, 0, 0, 0, 0, 0, 0) \\ \hat{\mathbf{y}}^{(3)} &= (\hat{y}_1^{(3)}, \hat{y}_2^{(3)}, \hat{y}_3^{(3)}, \hat{y}_4^{(3)}, \hat{y}_5^{(3)}, \hat{y}_6^{(3)}, \hat{y}_7^{(3)}, \hat{y}_8^{(3)}) = (0, 0, 1, 0, 0, 0, 0, 0) \\ \hat{\mathbf{y}}^{(4)} &= (\hat{y}_1^{(4)}, \hat{y}_2^{(4)}, \hat{y}_3^{(4)}, \hat{y}_4^{(4)}, \hat{y}_5^{(4)}, \hat{y}_6^{(4)}, \hat{y}_7^{(4)}, \hat{y}_8^{(4)}) = (0, 0, 0, 1, 0, 0, 0, 0) \\ \hat{\mathbf{y}}^{(5)} &= (\hat{y}_1^{(5)}, \hat{y}_2^{(5)}, \hat{y}_3^{(5)}, \hat{y}_4^{(5)}, \hat{y}_5^{(5)}, \hat{y}_6^{(5)}, \hat{y}_7^{(5)}, \hat{y}_8^{(5)}) = (0, 0, 0, 0, 1, 0, 0, 0) \\ \hat{\mathbf{y}}^{(6)} &= (\hat{y}_1^{(6)}, \hat{y}_2^{(6)}, \hat{y}_3^{(6)}, \hat{y}_4^{(6)}, \hat{y}_5^{(6)}, \hat{y}_6^{(6)}, \hat{y}_7^{(6)}, \hat{y}_8^{(6)}) = (0, 0, 0, 0, 0, 1, 0, 0) \\ \hat{\mathbf{y}}^{(7)} &= (\hat{y}_1^{(7)}, \hat{y}_2^{(7)}, \hat{y}_3^{(7)}, \hat{y}_4^{(7)}, \hat{y}_5^{(7)}, \hat{y}_6^{(7)}, \hat{y}_7^{(7)}, \hat{y}_8^{(7)}) = (0, 0, 0, 0, 0, 0, 1, 0) \\ \hat{\mathbf{y}}^{(8)} &= (\hat{y}_1^{(8)}, \hat{y}_2^{(8)}, \hat{y}_3^{(8)}, \hat{y}_4^{(8)}, \hat{y}_5^{(8)}, \hat{y}_6^{(8)}, \hat{y}_7^{(8)}, \hat{y}_8^{(8)}) = (0, 0, 0, 0, 0, 0, 0, 1) \end{aligned}$$

Mit Hilfe des MATLAB-Tools erhalten wir keine vollauf befriedigenden Ergebnisse, jeweils 4 davon sind erheblich fehlerbehaftet (rot eingezeichnet).

encoder_outputs =

0.9999	0.0000	0.0001	0.0001	0.0000	0.0000	0.0000	0.0000
0.0000	0.0001	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
0.0001	0.0148	0.9999	0.0001	0.0000	0.0002	0.0001	0.0450
0.0000	0.0001	0.0001	0.9999	0.0000	0.0000	0.0001	0.0000
0.0001	0.0000	0.0000	0.0001	0.9999	0.0001	0.0001	0.0000
0.0000	0.0001	0.0001	0.0000	0.0001	0.9999	0.0000	0.0022
0.0000	1.0000	0.0000	0.0001	0.0001	0.0000	0.9999	0.9739
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Mathematikaufgabe 95

Die Gewichte ergeben sich zu

```
[-0.52061 -0.1811 1.2399 -1.2388 -2.0993 0.95104 -0.3504 -0.059187;  
2.0707 -0.77878 0.65304 0.69502 -0.094855 -0.11686 -1.2808 -0.042635;  
1.4204 -0.81396 -0.29459 -0.52188 0.7182 1.6748 -0.35029 -0.21889]
```

Die Fehler sind erwartungsgemäß proportional zu den Netzantworten:

encoder_errors =

```
0.0001 -0.0000 -0.0001 -0.0001 -0.0000 -0.0000 -0.0000 -0.0000  
-0.0000 0.9999 -0.0001 -0.0000 -0.0000 -0.0000 -0.0000 -0.0000  
-0.0001 -0.0148 0.0001 -0.0001 -0.0000 -0.0002 -0.0001 -0.0450  
-0.0000 -0.0001 -0.0001 0.0001 -0.0000 -0.0000 -0.0001 -0.0000  
-0.0001 -0.0000 -0.0000 -0.0001 0.0001 -0.0001 -0.0001 -0.0000  
-0.0000 -0.0001 -0.0001 -0.0000 -0.0001 0.0001 -0.0000 -0.0022  
-0.0000 -1.0000 -0.0000 -0.0001 -0.0001 -0.0000 0.0001 -0.9739  
-0.0000 -0.0000 -0.0000 -0.0000 -0.0000 -0.0000 -0.0000 1.0000
```

Grund dafür ist, daß unser Lösungsvektor nicht das globale Minimum gefunden hat, sondern in ein Seitenminimum gelaufen ist (auch Tools haben ihre Schwächen). Solche Probleme abzustellen ist zeitaufwendig und soll daher dem Leser überlassen bleiben.