

Training eines Perzeptrons am Beispiel eines logischen Oders

Das logische Oder wird bereits durch die Angabe der folgenden vier Trainingsdatensätze komplett definiert. Die Bias-Einheit x_0 ist immer 1.

x_0	x_1	x_2	y
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

Das Training erfolgt nach dem folgenden Algorithmus:

1. Die Gewichte werden mit 0 oder zufälligen Werten initialisiert.
2. Die Trainingsdatensätze werden nacheinander in das Perzeptron eingesetzt und der Output berechnet. Dann werden die Gewichte gemäß der Formel
$$w_i^{\text{neu}} = w_i^{\text{alt}} + \alpha (y - \hat{y}^{\text{alt}}) x_i$$
für $i = 0 \dots m$ aktualisiert.
3. Schritt 2 wird solange wiederholt, bis sich die Gewichte nicht mehr ändern und trotzdem für alle Trainingsdatensätze die richtige Klasse prognostiziert wird.

① Wir initialisieren die Gewichte mit $\vec{w}^T = (0, 0, 0)$ und wählen $\alpha = 1$.

② 1. Trainingsdatensatz:

$$\hat{y}^{\text{alt}} = \phi\left((0, 0, 0) \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) = \phi(0) = 0, \text{ aber } y = 1$$

Aktualisierung Gewichte:

$$\begin{aligned} w_0 &= 0 + 1 \cdot (1 - 0) \cdot 1 = 1 \\ w_1 &= 0 + 1 \cdot (1 - 0) \cdot 0 = 0 \\ w_2 &= 0 + 1 \cdot (1 - 0) \cdot 0 = 0 \end{aligned}$$

2. Trainingsdatensatz:

$$\hat{y}^{att} = \phi((-1, 0, 0) \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}) = \phi(-1) = 0, \text{ aber } y = 1$$



Aktualisierung Gewichte:

$$\begin{aligned} \omega_0 &= -1 + 1 \cdot (1-0) \cdot 1 = 0 \\ \omega_1 &= 0 + 1 \cdot (1-0) \cdot 0 = 0 \\ \omega_2 &= 0 + 1 \cdot (1-0) \cdot 1 = 1 \end{aligned}$$

3. Trainingsdatensatz

$$\hat{y}^{att} = \phi(0, 0, 1) \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \phi(0) = 1 \text{ und } y = 1$$



keine Aktualisierung der Gewichte notwendig

4. Trainingsdatensatz

$$\hat{y}^{att} = \phi(0, 0, 1) \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \phi(1) = 1 \text{ und } y = 1$$



keine Aktualisierung der Gewichte notwendig

1. Trainingsdatensatz

$$\hat{y}^{att} = \phi(0, 0, 1) \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \phi(0) = 1, \text{ aber } y = 0$$



Aktualisierung Gewichte:

$$\begin{aligned} \omega_0 &= 0 + 1 \cdot (0-1) \cdot 1 = -1 \\ \omega_1 &= 0 + 1 \cdot (0-1) \cdot 0 = 0 \\ \omega_2 &= 1 + 1 \cdot (0-1) \cdot 0 = 1 \end{aligned}$$

2. Trainingsdatensatz:

$$\hat{y}^{att} = \phi(-1, 0, 1) \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \phi(1) = 1 \text{ und } y = 1$$



keine Aktualisierung der Gewichte notwendig

3. Trainingsdatensatz

$$\hat{y}^{\text{alt}} = \phi((-1, 0, 1) \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}) = \phi(-1) = 0, \text{ aber } y=1 \quad \text{⚡}$$

Aktualisierung Gewichte:

$$\begin{aligned} \omega_0 &= -1 + 1 \cdot (1-0) \cdot 1 = 0 \\ \omega_1 &= 0 + 1 \cdot (1-0) \cdot 1 = 1 \\ \omega_2 &= 1 + 1 \cdot (1-0) \cdot 0 = 1 \end{aligned}$$

4. Trainingsdatensatz

$$\hat{y}^{\text{alt}} = \phi(0, 1, 1) \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \phi(2) = 1 \text{ und } y=1 \quad \checkmark$$

keine Aktualisierung der Gewichte notwendig

1. Trainingsdatensatz

$$\hat{y}^{\text{alt}} = \phi(0, 1, 1) \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \phi(0) = 1, \text{ aber } y=0 \quad \text{⚡}$$

Aktualisierung Gewichte:

$$\begin{aligned} \omega_0 &= 0 + 1 \cdot (0-1) \cdot 1 = -1 \\ \omega_1 &= 1 + 1 \cdot (0-1) \cdot 0 = 1 \\ \omega_2 &= 1 + 1 \cdot (0-1) \cdot 0 = 1 \end{aligned}$$

2. Trainingsdatensatz

$$\hat{y}^{\text{alt}} = \phi(-1, 1, 1) \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \phi(0) = 1 \text{ und } y=1 \quad \checkmark$$

keine Aktualisierung der Gewichte notwendig

3. Trainingsdatensatz

$$\hat{y}^{\text{alt}} = \phi(-1, 1, 1) \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \phi(0) = 1 \text{ und } y=1 \quad \checkmark$$

keine Aktualisierung der Gewichte notwendig

4. Trainingsdatensatz:

$$\hat{y}^{\text{alt}} = \phi(-1, 1, 1) \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \phi(1) = 1 \text{ und } y=1 \quad \checkmark$$

keine Aktualisierung der Gewichte notwendig

1. Trainingsdatensatz :

$$\hat{y}^{\text{akt}} = \phi \left((-1, 1, 1) \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) = \phi(-1) = 0 \quad \text{und} \quad y = 1 \quad \checkmark$$

keine Aktualisierung der Gewichte notwendig

③ Bei den letzten vier (und damit allen) Trainingsdatensätzen war keine Aktualisierung der Gewichte mehr notwendig. Wir sind mit dem Training des Perzeptrons fertig. Die Lernregel für das logische Oder lautet also

$$\vec{\omega}^T = (-1, 1, 1) \quad \text{mit Aktivierungsfunktion} \quad \phi(x) = \begin{cases} 0, & x < 0 \\ 1 & x \geq 0 \end{cases}.$$