

PROSEMINAR: DATA MINING

Neuronale Netze: Grundlagen

Lukas Krenz

12. Juni 2015

Technische Universität München

Neuronale Netze sind ein:

- graphisches
- nicht-lineares
- flexibles

Modell für maschinelles Lernen.

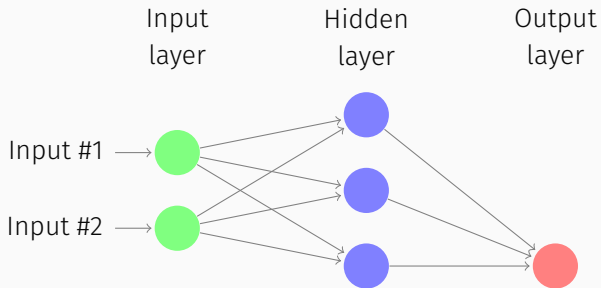


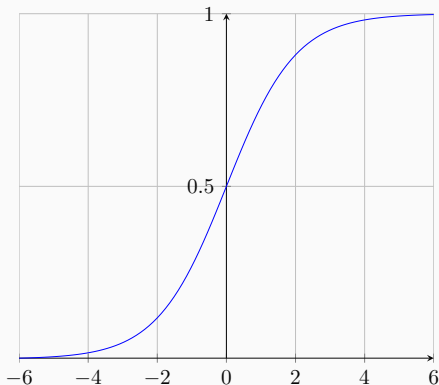
Abbildung 1: Ein 2-schichtiges MLP.

$$\sigma_1(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

$$\sigma_2(x) = \tanh(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}} \quad (2)$$

$$\sigma_3(x) = \max(0, x) \quad (3)$$

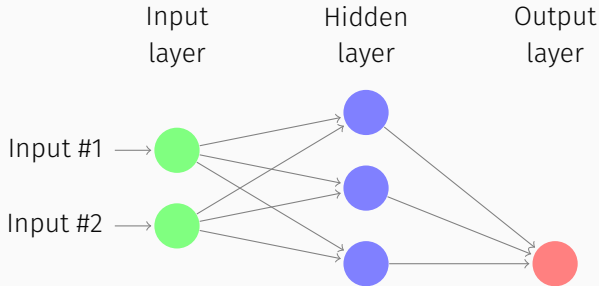
Abbildung 2: Die logistische Funktion



$$\sigma_1(x) = \frac{1}{1 + e^{-x}} \quad (4)$$

$$\sigma_2(x) = \tanh(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}} \quad (5)$$

$$\sigma_3(x) = \max(0, x) \quad (6)$$



$$\text{net}_j = \sum_i w_{ji} y_i = w_j^t x \quad (7)$$

$$y_j = \sigma(\text{net}_j) \quad (8)$$

$$E = \sum_{k=1}^K \sum_{i=1}^N \left(y_i^k - t_i^k \right)^2 \quad (9)$$

$$= \sum_n E_n(y_1, \dots, y_c) \quad (10)$$

$$\frac{\partial E^n}{\partial w_{ij}} = \frac{\partial E^n}{\partial net_j} \frac{\partial net_j}{\partial w_{ji}} \quad (11)$$

$$\frac{\partial net_j}{\partial w_{ji}} = y_i \quad (12)$$

$$\frac{\partial E^n}{\partial net_j} = \delta_j \quad (13)$$

$$\frac{\partial E^n}{\partial w_{i,j}} = y_i \delta_j. \quad (14)$$

$$\delta_j = \begin{cases} \sigma'(net_j)(y_j - \hat{y}_j) & \text{wenn } j \text{ Ausgabeneuron ist} \\ \sigma'(net_j) \sum_k w_{kj} \delta_k & \text{wenn } j \text{ verdecktes Neuron ist} \end{cases} \quad (15)$$

$$x_{n+1} = x_n - \eta \nabla F(x_n) \quad (16)$$

$$\Delta_w^t = -\eta \nabla_w E(w) \quad (17)$$

Vorteile:

- mächtig
- flexibel

Nachteile

- viele Parameter, kompliziert anzupassen
- langsam

In der **aktuellen** Forschung relevant: Netze anderer Topologien