Radiale-Basisfunktionen-Netze

Eigenschaften von Radiale-Basisfunktionen-Netzen (RBF-Netzen)

- RBF-Netze sind streng geschichtete, vorwärtsbetriebene neuronale Netze mit genau einer versteckten Schicht.
- Als Netzeingabe- und Aktivierungsfunktion werden radiale Basisfunktionen verwendet.
- Jedes Neuron erhält eine Art "Einzugsgebiet".
- Die Gewichte der Verbindungen von der Eingabeschicht zu einem Neuron geben das Zentrum an.

Ein **radiale-Basisfunktionen-Netz (RBF-Netz)** ist ein neuronales Netz mit einem Graph $G = (U, C)$, das die folgenden Bedingungen erfüllt:

(i) $U_{\text{in}} \cap U_{\text{out}} = \emptyset$,

(ii) $C = (U_{\text{in}} \times U_{\text{hidden}}) \cup C', \quad C' \subseteq (U_{\text{hidden}} \times U_{\text{out}})$

Die Netzeingabefunktion jedes versteckten Neurons ist eine **Abstandsfunktion** zwischen dem Eingabevektor und dem Gewichtsvektor, d.h.

$$\forall u \in U_{\text{hidden}} : \qquad f_{\text{net}}^{(u)}(\boldsymbol{w}_u, \mathbf{in}_u) = d(\boldsymbol{w}_u, \mathbf{in}_u),$$

wobei $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_0^+$ eine Funktion ist, die $\forall \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n :$ erfüllt:

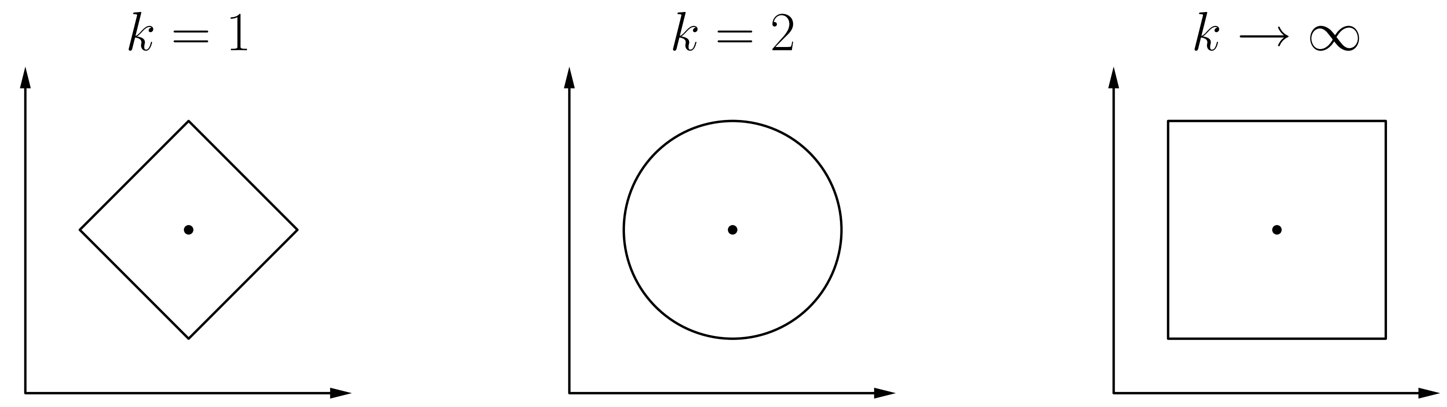
$$\begin{array}{rlr} (i) & d(\boldsymbol{x}, \boldsymbol{y}) = 0 \quad \Leftrightarrow \quad \boldsymbol{x} = \boldsymbol{y}, & \\ (ii) & d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{y}, \boldsymbol{x}) & \text{(Symmetrie)}, \\ (iii) & d(\boldsymbol{x}, \boldsymbol{z}) \leq d(\boldsymbol{x}, \boldsymbol{y}) + d(\boldsymbol{y}, \boldsymbol{z}) & \text{(Dreiecksungleichung)}. \end{array}$$

Veranschaulichung von Abstandsfunktionen

$$d_k(\boldsymbol{x}, \boldsymbol{y}) = \left(\sum_{i=1}^{n} (x_i - y_i)^k\right)^{\frac{1}{k}}$$

Bekannte Spezialfälle dieser Familie sind:

$k = 1$: Manhattan-Abstand ,
$k = 2$: Euklidischer Abstand,
$k \to \infty$: Maximum-Abstand, d.h. $d_\infty(\boldsymbol{x}, \boldsymbol{y}) = \max_{i=1}^{n} |x_i - y_i|$.



(alle Punkte auf dem Kreis bzw. den Vierecken haben denselben Abstand zum Mittelpunkt, entsprechend der jeweiligen Abstandsfunktion)

Radiale-Basisfunktionen-Netze

Die Netzeingabefunktion der Ausgabeneuronen ist die gewichtete Summe ihrer Eingaben, d.h.

$$\forall u \in U_{\text{out}}: \qquad f_{\text{net}}^{(u)}(\boldsymbol{w}_u, \mathbf{in}_u) = \boldsymbol{w}_u \mathbf{in}_u = \sum_{v \in \text{pred}\,(u)} w_{uv} \,\text{out}_v\,.$$

Die Aktivierungsfunktion jedes versteckten Neurons ist eine sogenannte **radiale Funktion**, d.h. eine monoton fallende Funktion

$$f: \mathbb{R}_0^+ \to [0,1] \quad \text{with} \quad f(0) = 1 \quad \text{and} \quad \lim_{x \to \infty} f(x) = 0.$$

Die Aktivierungsfunktion jedes Ausgabeneurons ist eine lineare Funktion

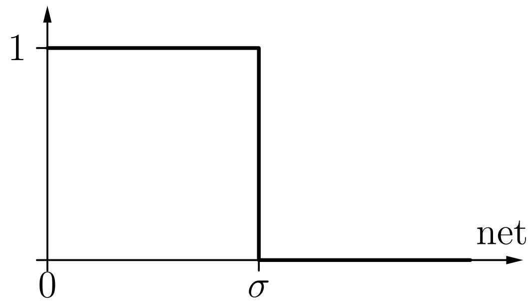
$$f_{\text{act}}^{(u)}(\text{net}_u, \theta_u) = \text{net}_u - \theta_u.$$

(Die lineare Aktivierungsfunktion ist wichtig für die Initialisierung.)

Radiale Aktivierungsfunktionen

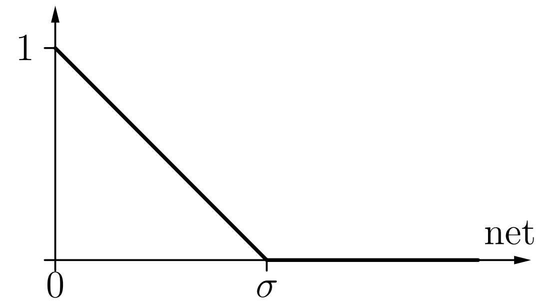
Rechteckfunktion:

$$f_{\text{act}}(\text{net}, \sigma) = \begin{cases} 0, & \text{falls net} > \sigma, \\ 1, & \text{sonst.} \end{cases}$$



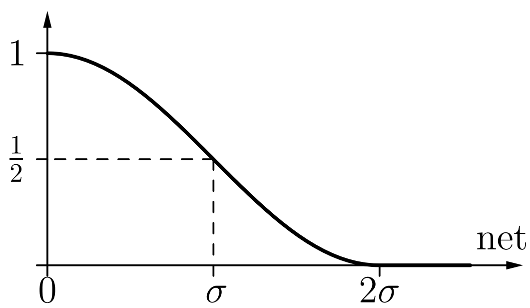
Dreiecksfunktion:

$$f_{\text{act}}(\text{net}, \sigma) = \begin{cases} 0, & \text{falls net} > \sigma, \\ 1 - \frac{\text{net}}{\sigma}, & \text{sonst.} \end{cases}$$



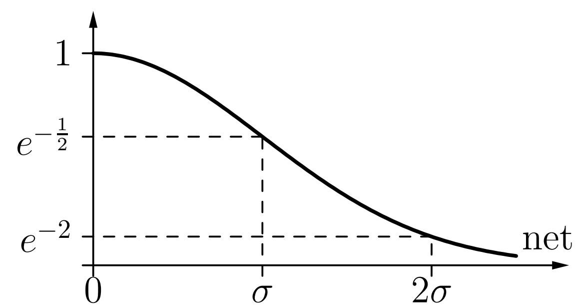
Kosinus bis Null:

$$f_{\text{act}}(\text{net}, \sigma) = \begin{cases} 0, & \text{falls net} > 2\sigma, \\ \frac{\cos\left(\frac{\pi}{2\sigma}\text{net}\right)+1}{2}, & \text{sonst.} \end{cases}$$

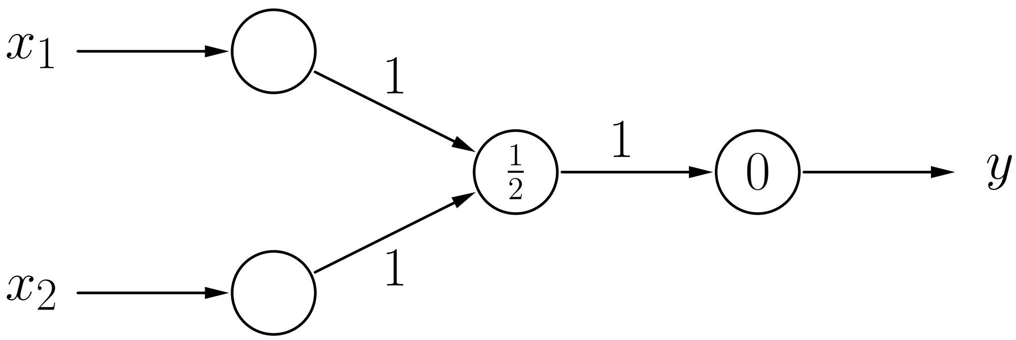


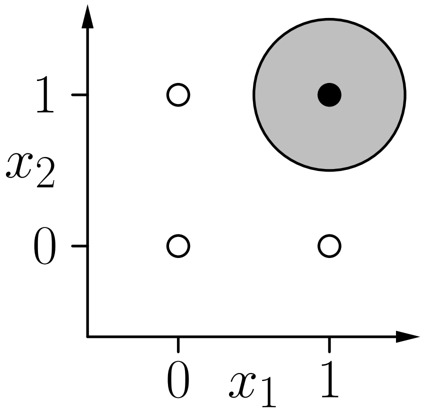
Gaußsche Funktion:

$$f_{\text{act}}(\text{net}, \sigma) = e^{-\frac{\text{net}^2}{2\sigma^2}}$$



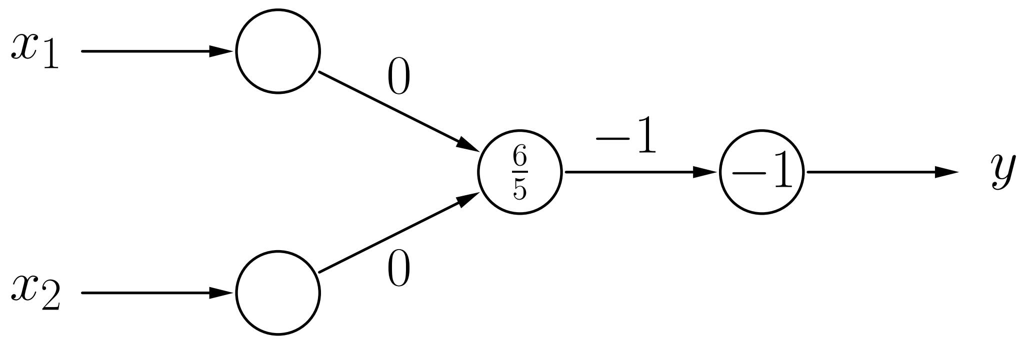
Radiale-Basisfunktionen-Netz für die Konjunktion $x_1 \wedge x_2$

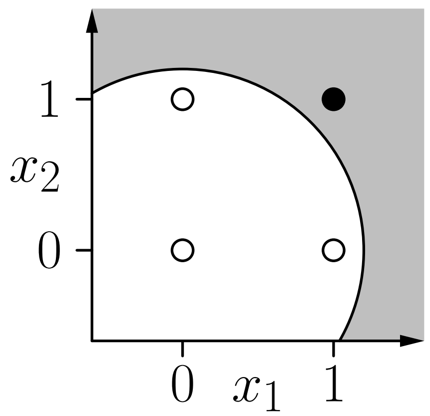




- (1,1) ist Zentrum
- Referenzradius ist $\frac{1}{2}$
- Euklidischer Abstand
- Rechteckfunktion als Aktivierung
- Biaswert 0 im Ausgabeneuron

Radiale-Basisfunktionen-Netz für die Konjunktion $x_1 \wedge x_2$



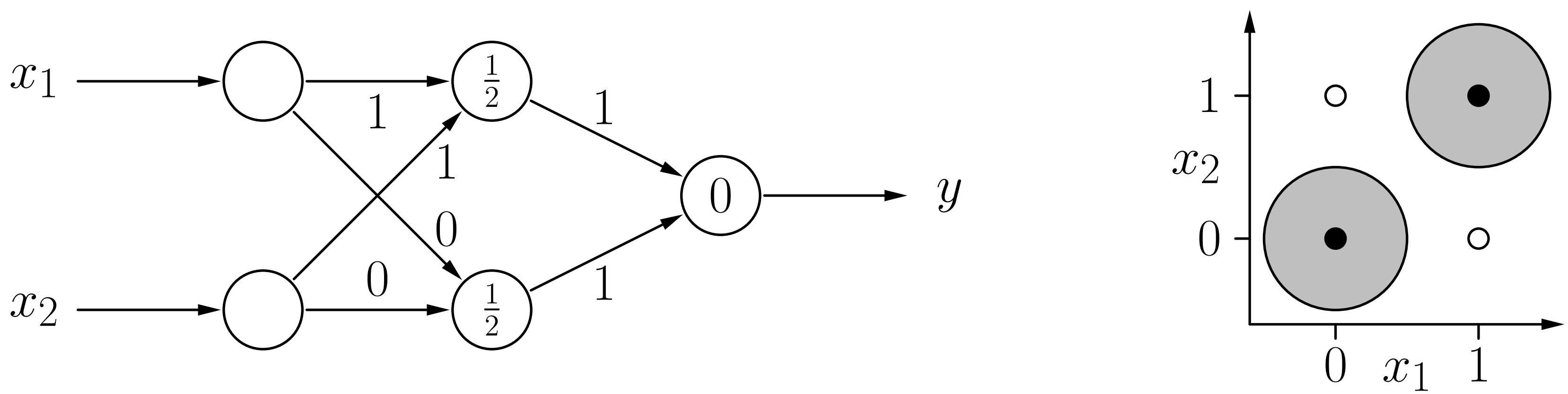


- (0,0) ist Zentrum
- Referenzradius ist $\frac{6}{5}$
- Euklidischer Abstand
- Rechteckfunktion als Aktivierung
- Biaswert -1 im Ausgabeneuron

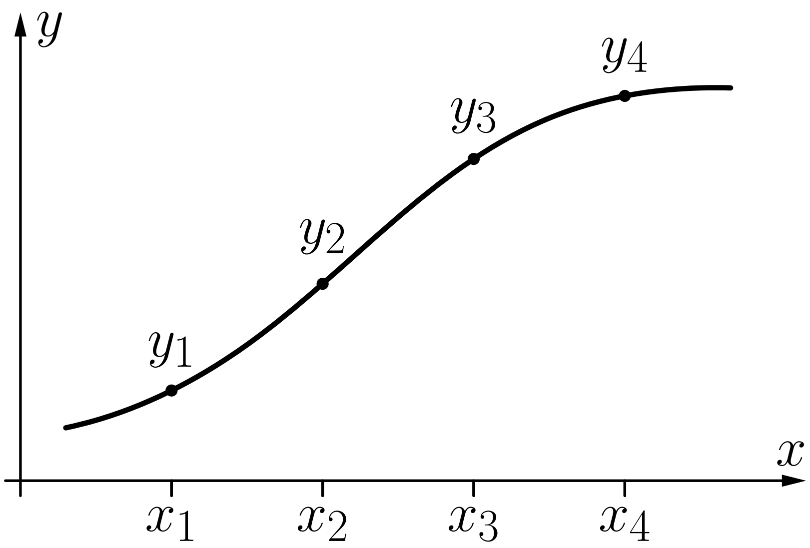
Radiale-Basisfunktionen-Netz für die Biimplikation $x_1 \leftrightarrow x_2$

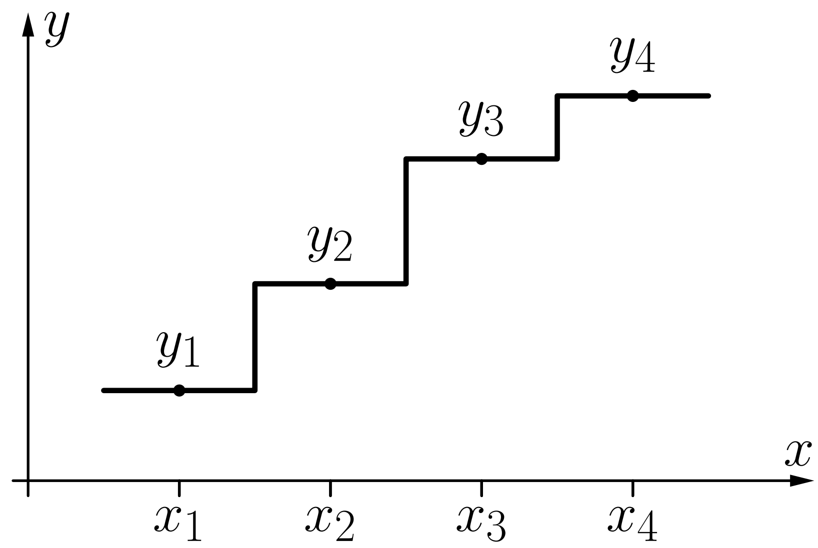
Idee: logische Zerlegung

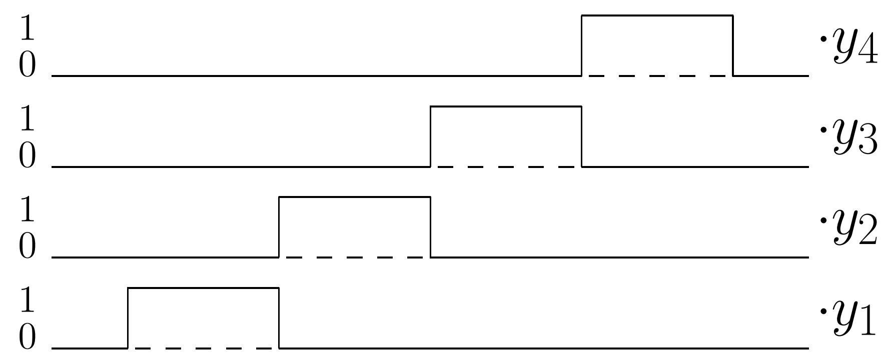
$$x_1 \leftrightarrow x_2 \quad \equiv \quad (x_1 \wedge x_2) \vee \neg(x_1 \vee x_2)$$



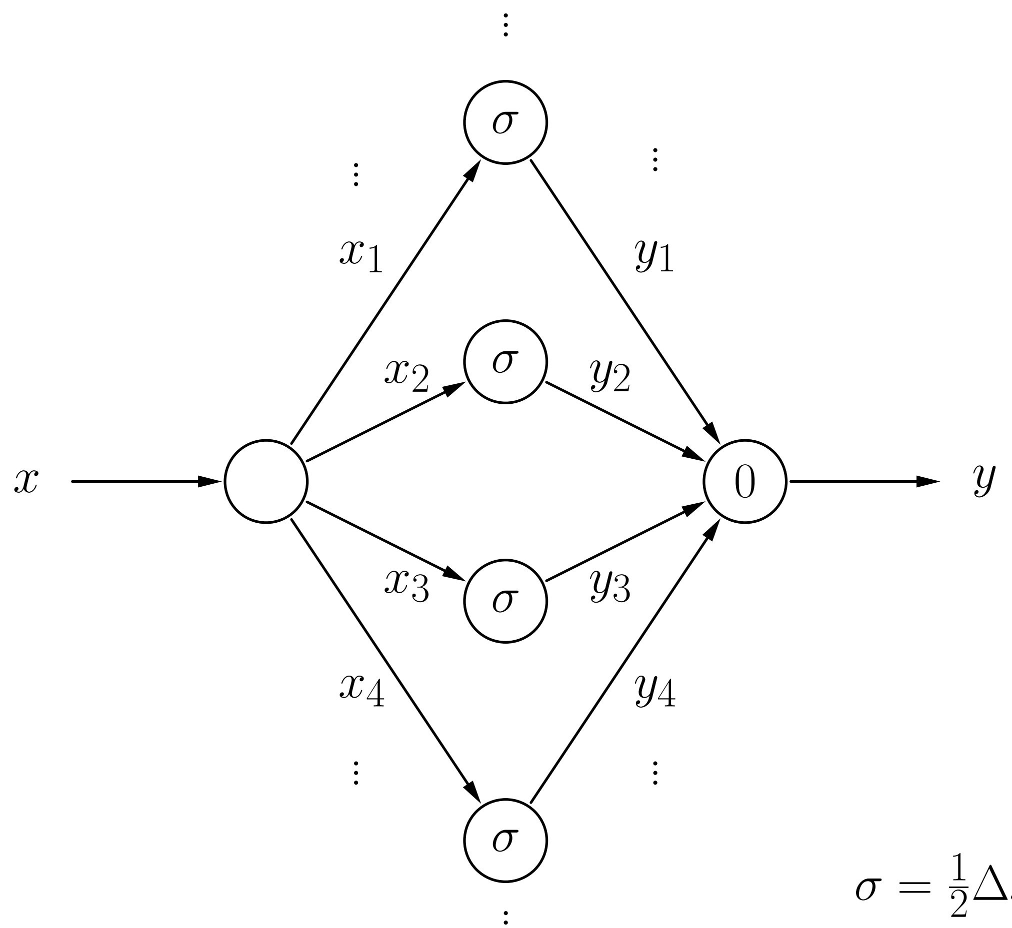
Radiale-Basisfunktionen-Netze: Funktionsapproximation





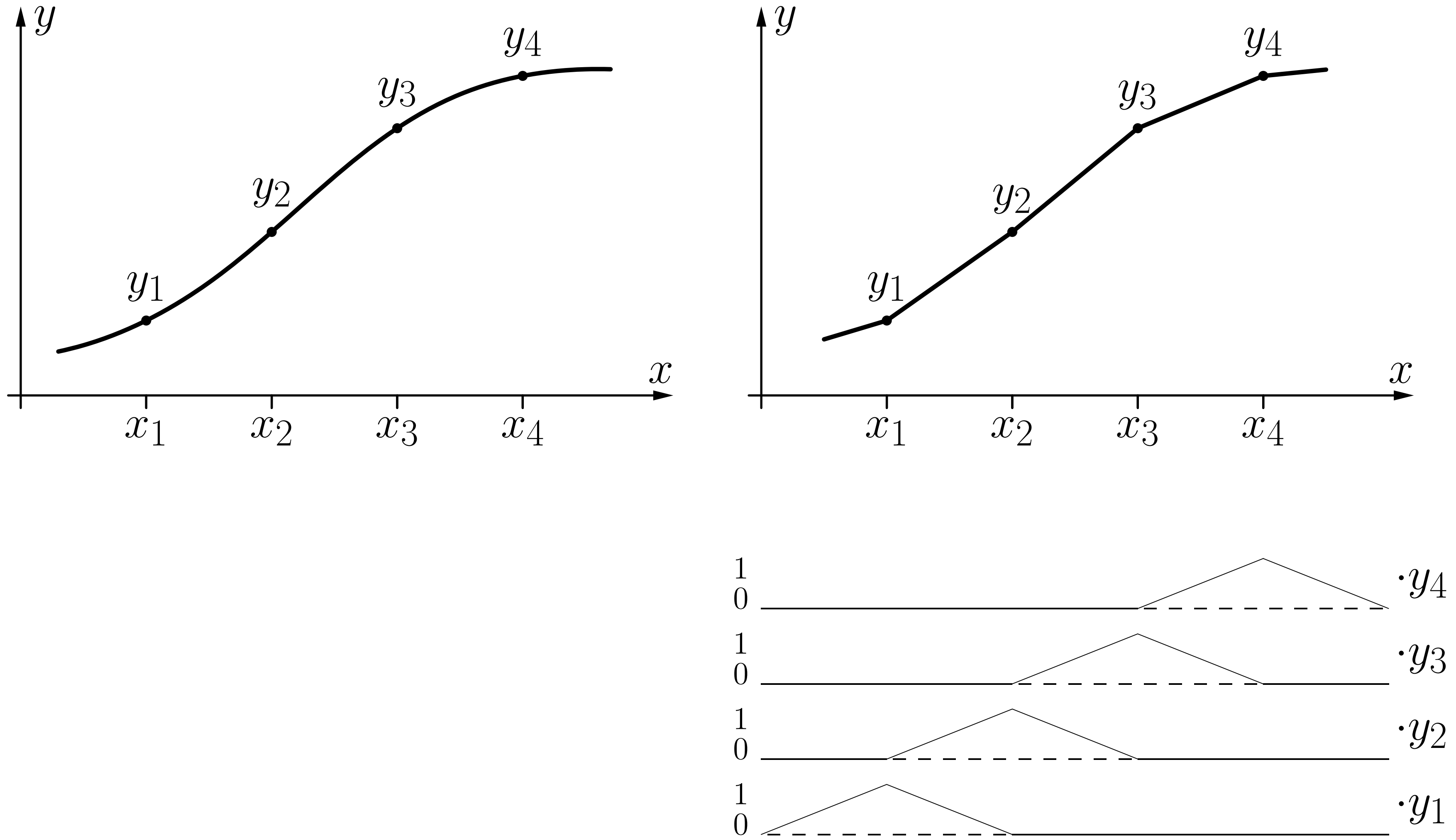


Annäherung der Originalfunktion durch Stufenfunktionen, deren Stufen durch einzelne Neuronen eines RBF-Netzes dargestellt werden können (vgl. MLPs).

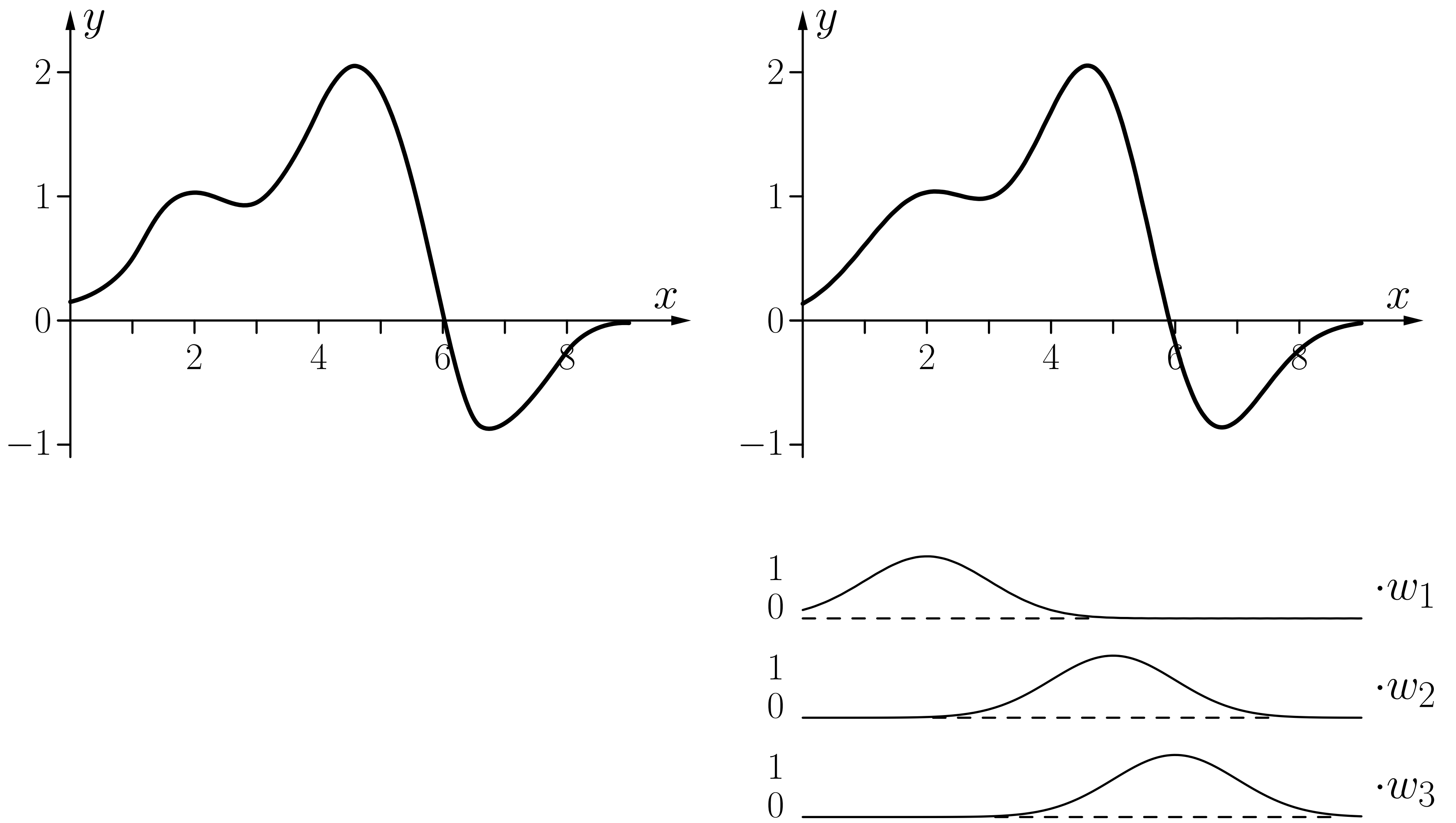


$$\sigma = \frac{1}{2}\Delta x = \frac{1}{2}(x_{i+1} - x_i)$$

Ein RBF-Netz, das die Treppenfunktion von der vorherigen Folie bzw. die stückweise lineare Funktion der folgenden Folie berechnet (dazu muss nur die Aktivierungsfunktion der versteckten Neuronen geändert werden).



Darstellung einer stückweise linearen Funktion durch eine gewichtete Summe von Dreiecksfunktionen mit Zentren x_i.



Annäherung einer Funktion durch eine Summe von Gaußkurven mit Radius $\sigma = 1$.
Es ist $w_1 = 2$, $w_2 = 3$ und $w_3 = -2$.

RBF-Netz für eine Summe dreier Gaußfunktionen