

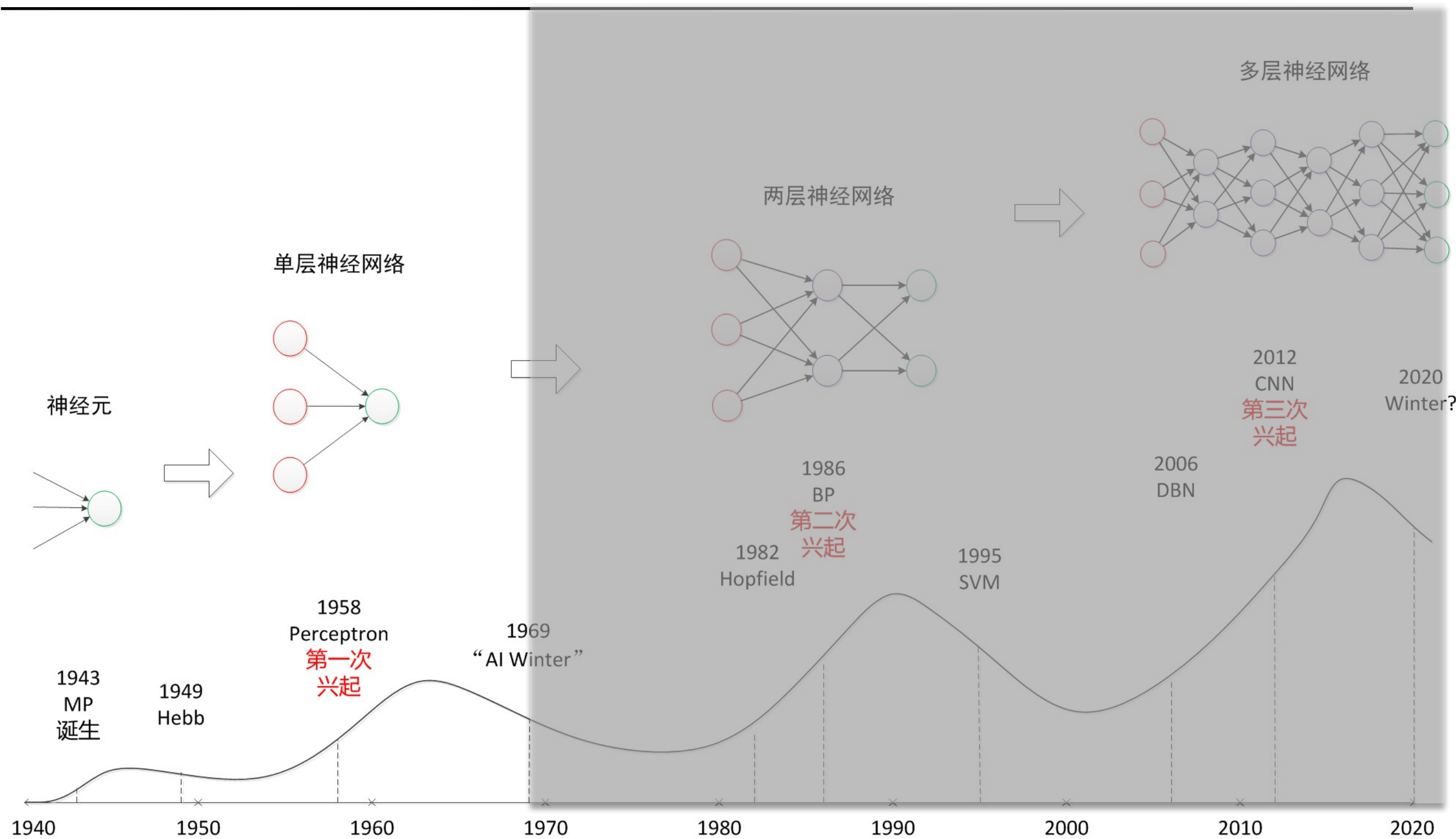
ThoughtWorks®

深度学习工作坊

神经网络介绍

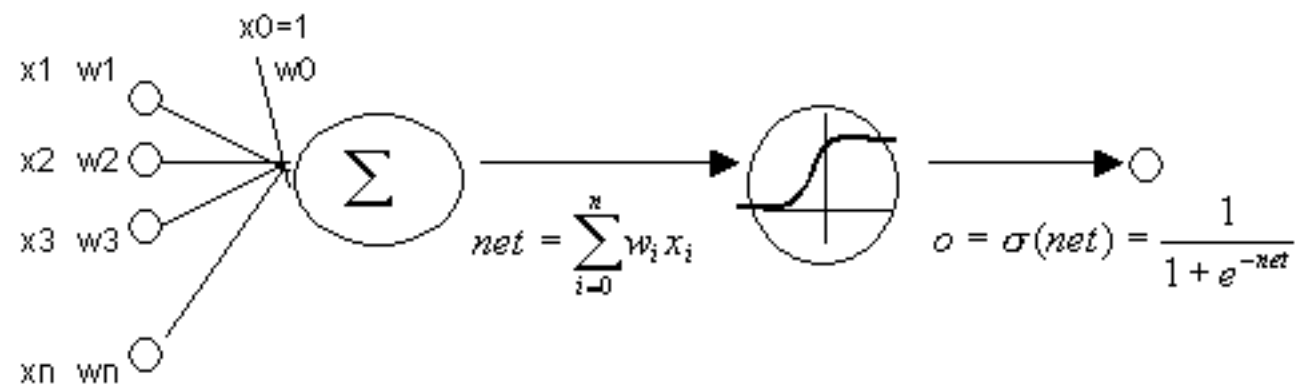
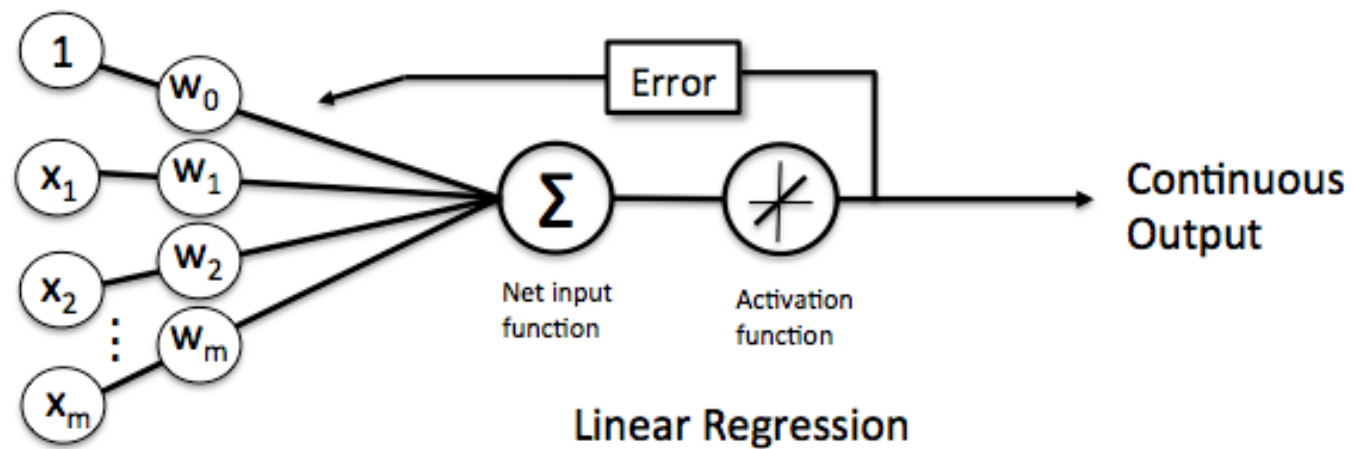
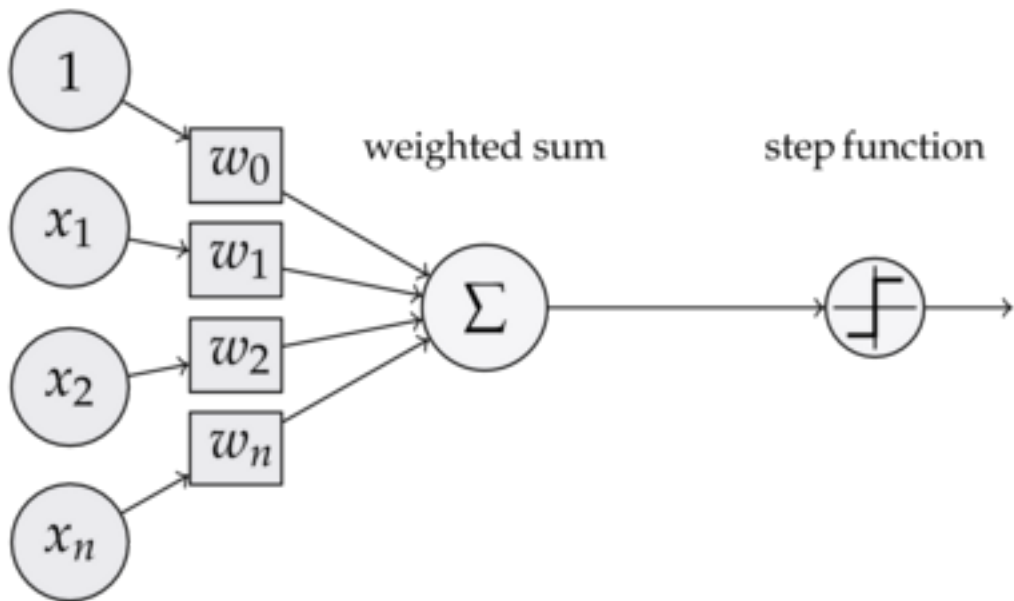
吴志平

神经网络的历史发展

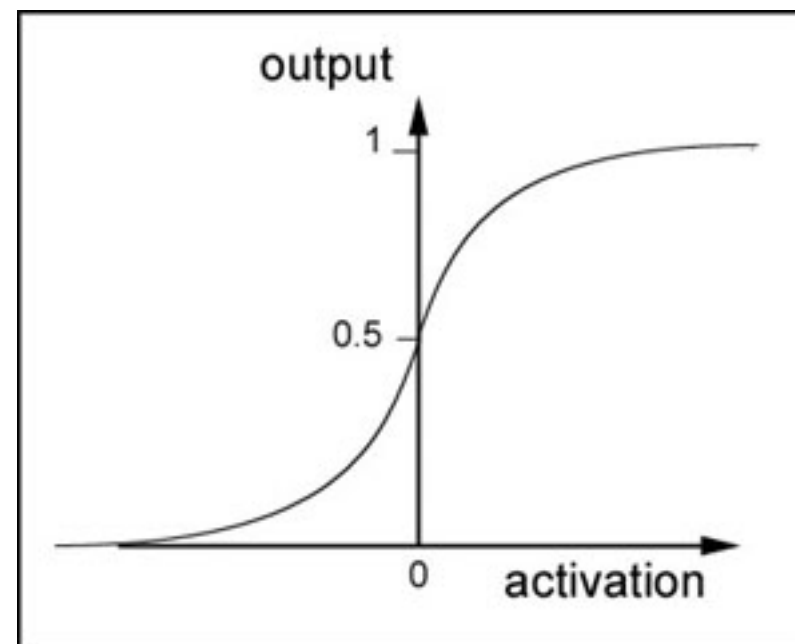


神经元

inputs weights



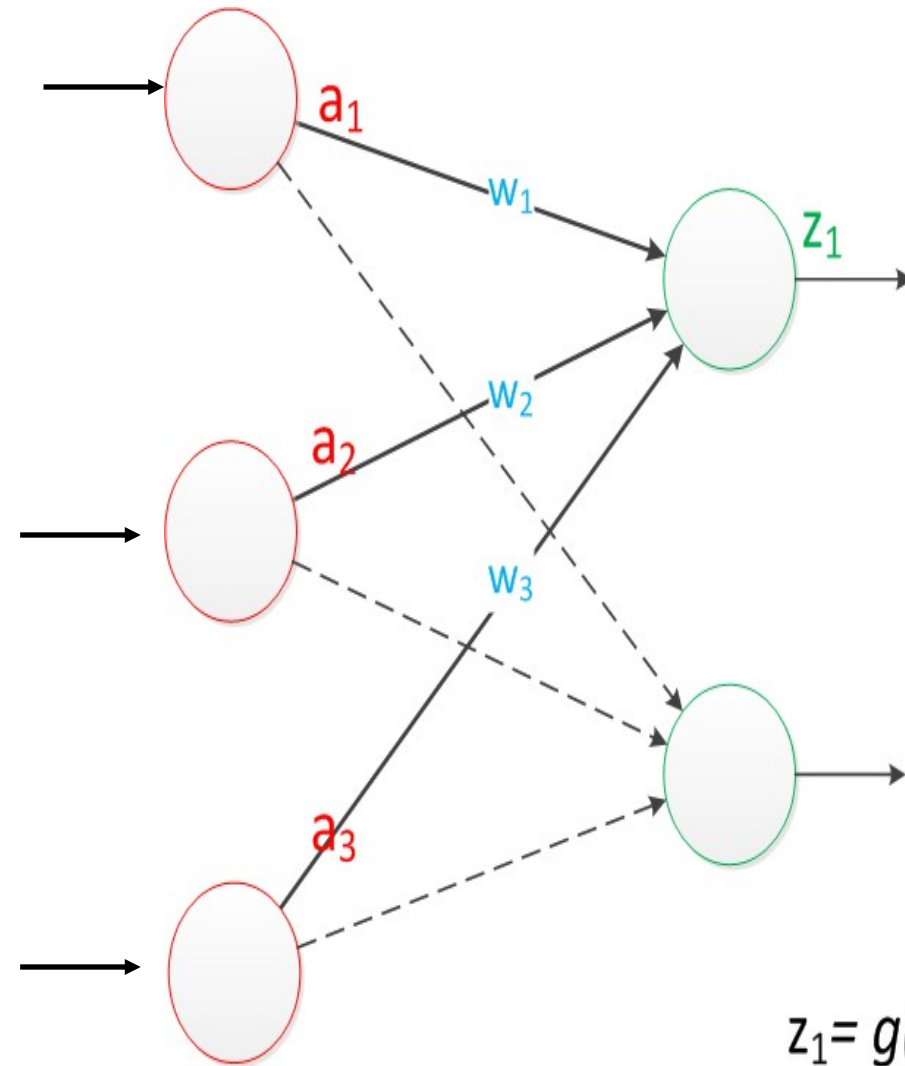
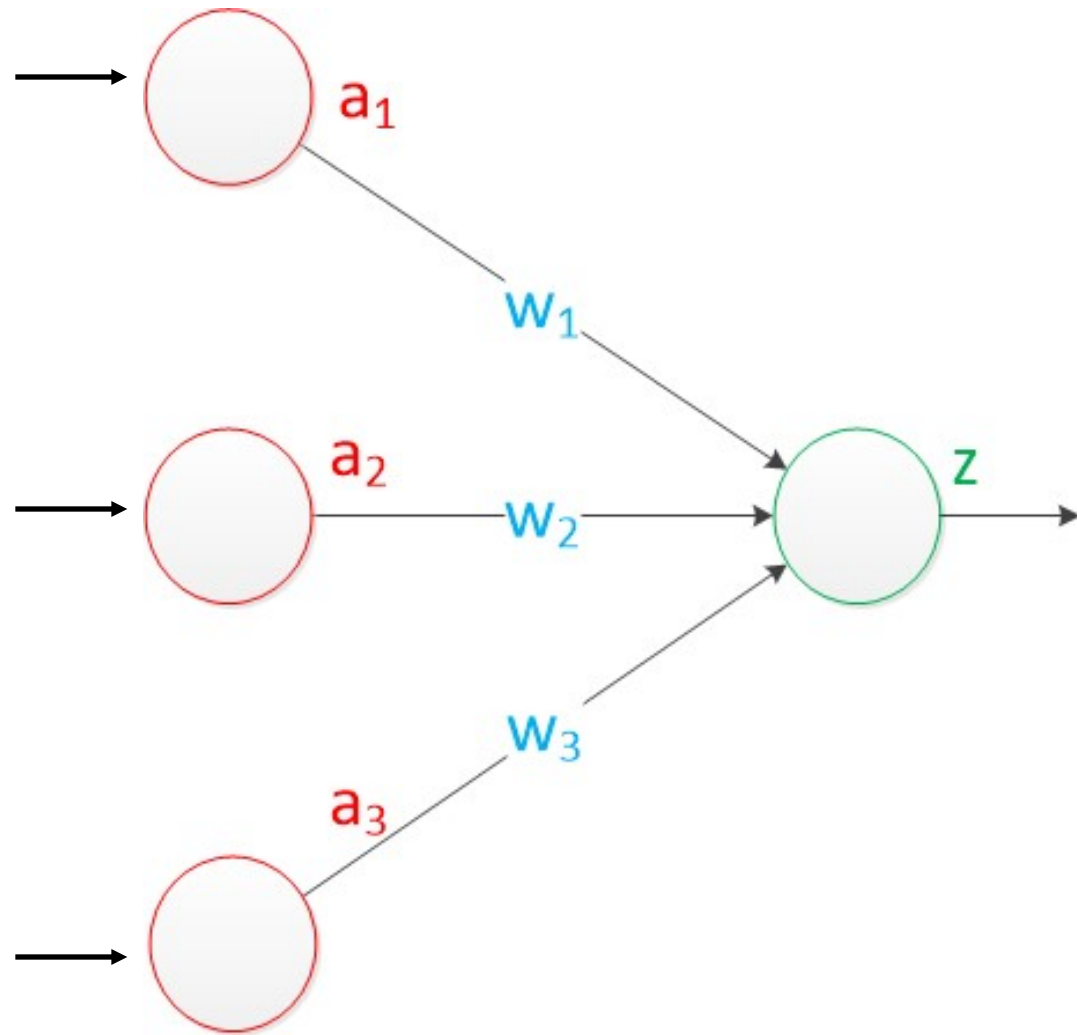
[The sigmoid threshold unit]



令 $y = \text{sigmoid}(x)$

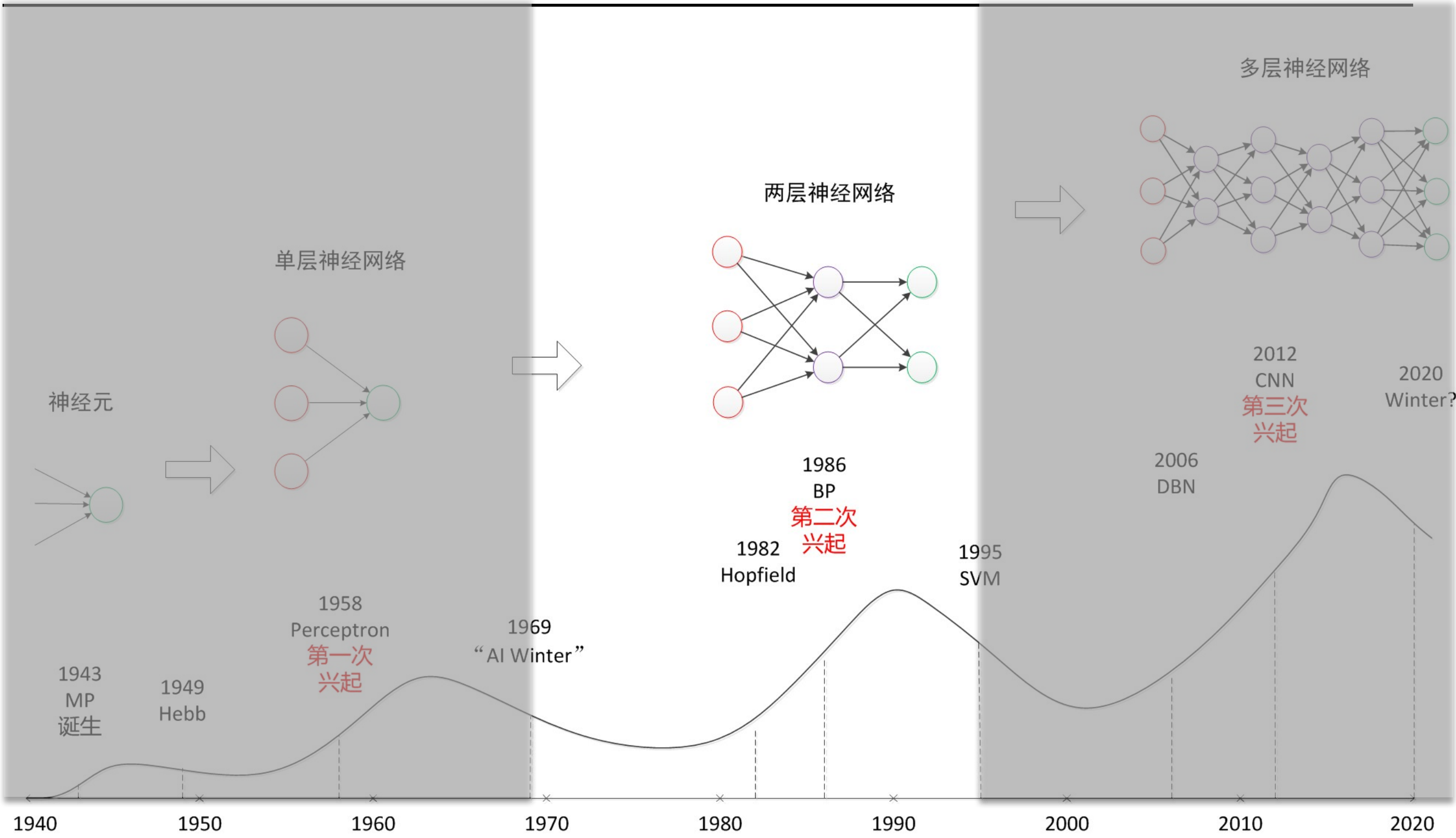
则 $y' = y(1 - y)$

单层神经网络

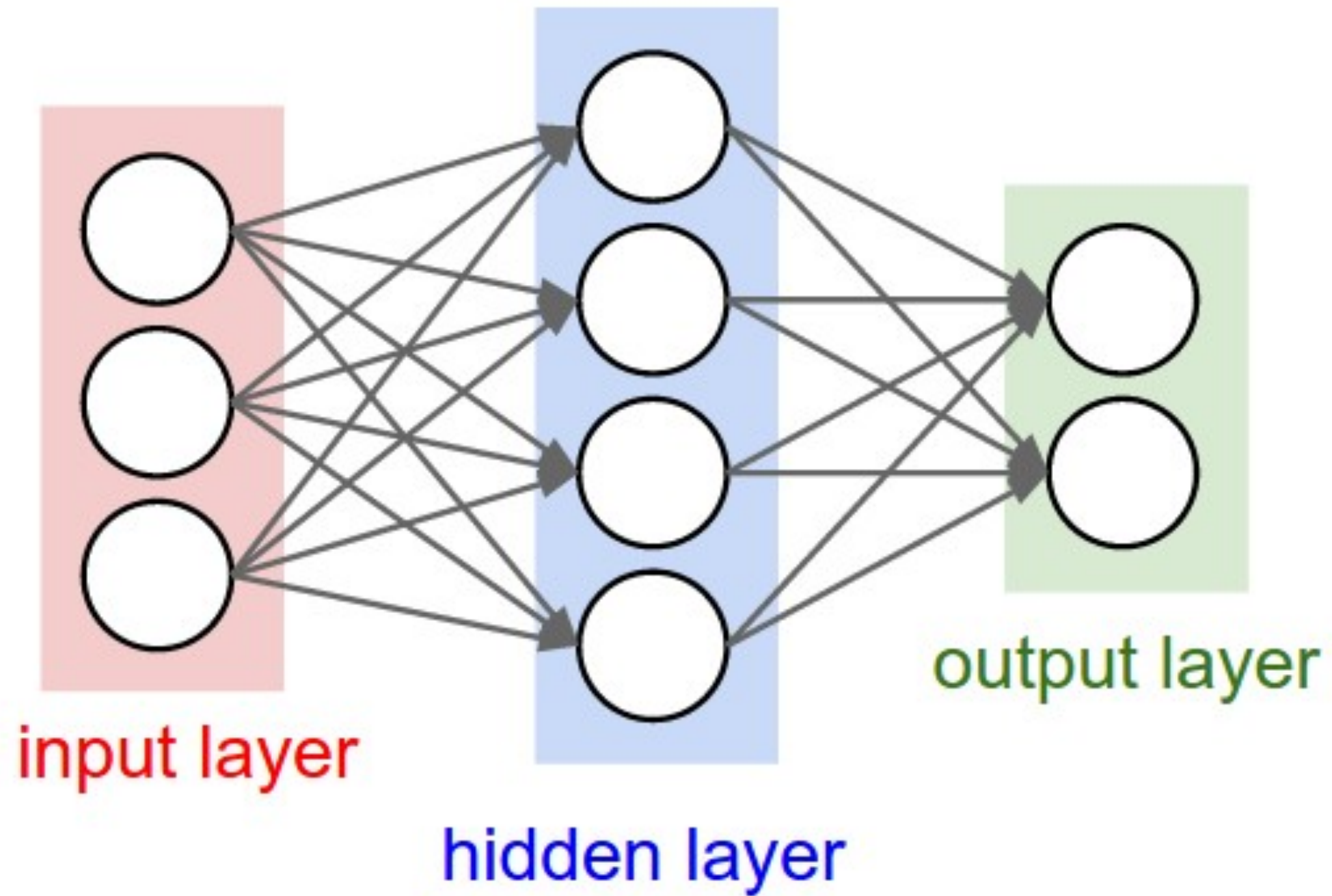


$$z_1 = g(a_1 * w_1 + a_2 * w_2 + a_3 * w_3)$$

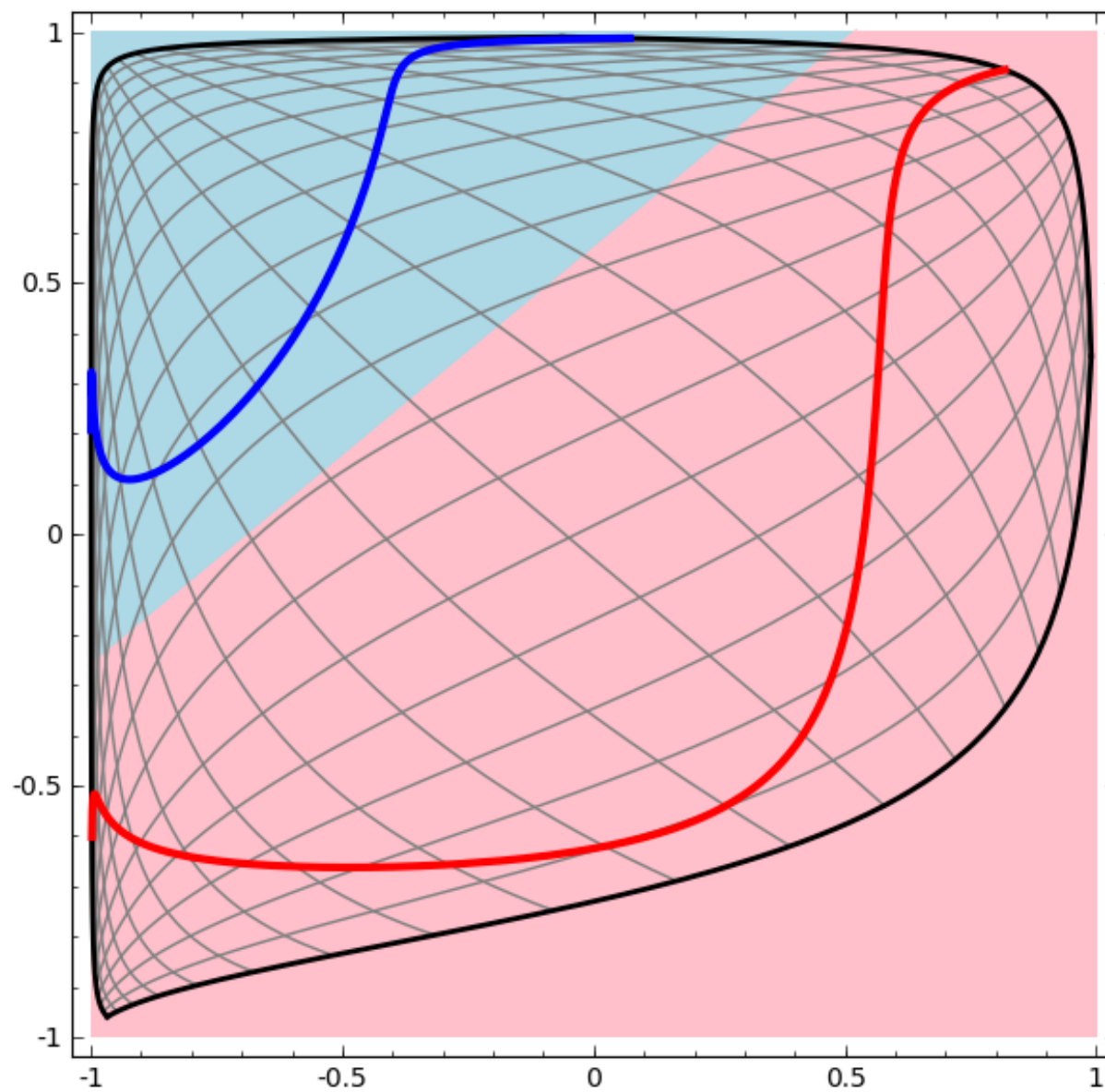
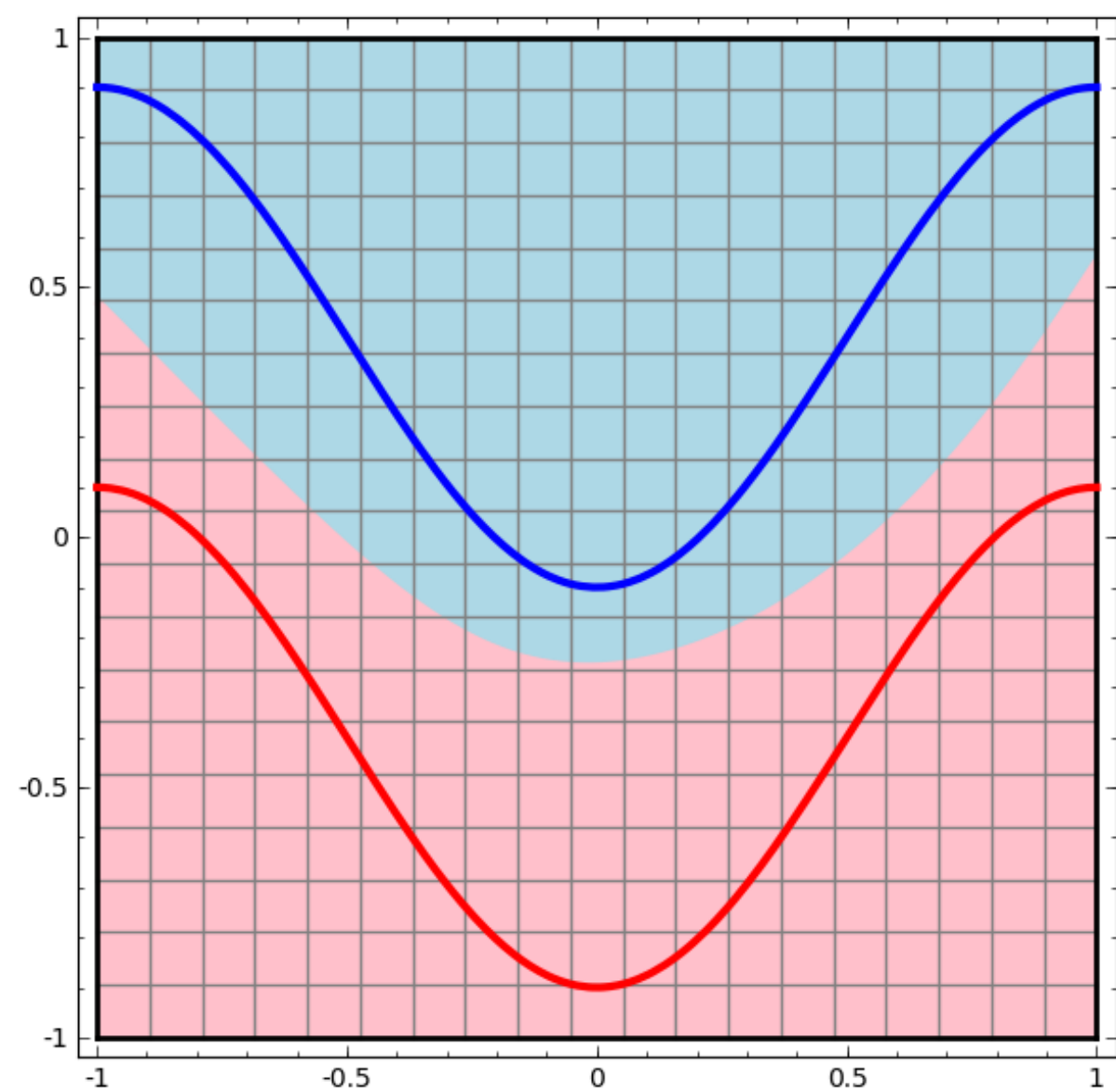
神经网络的历史发展



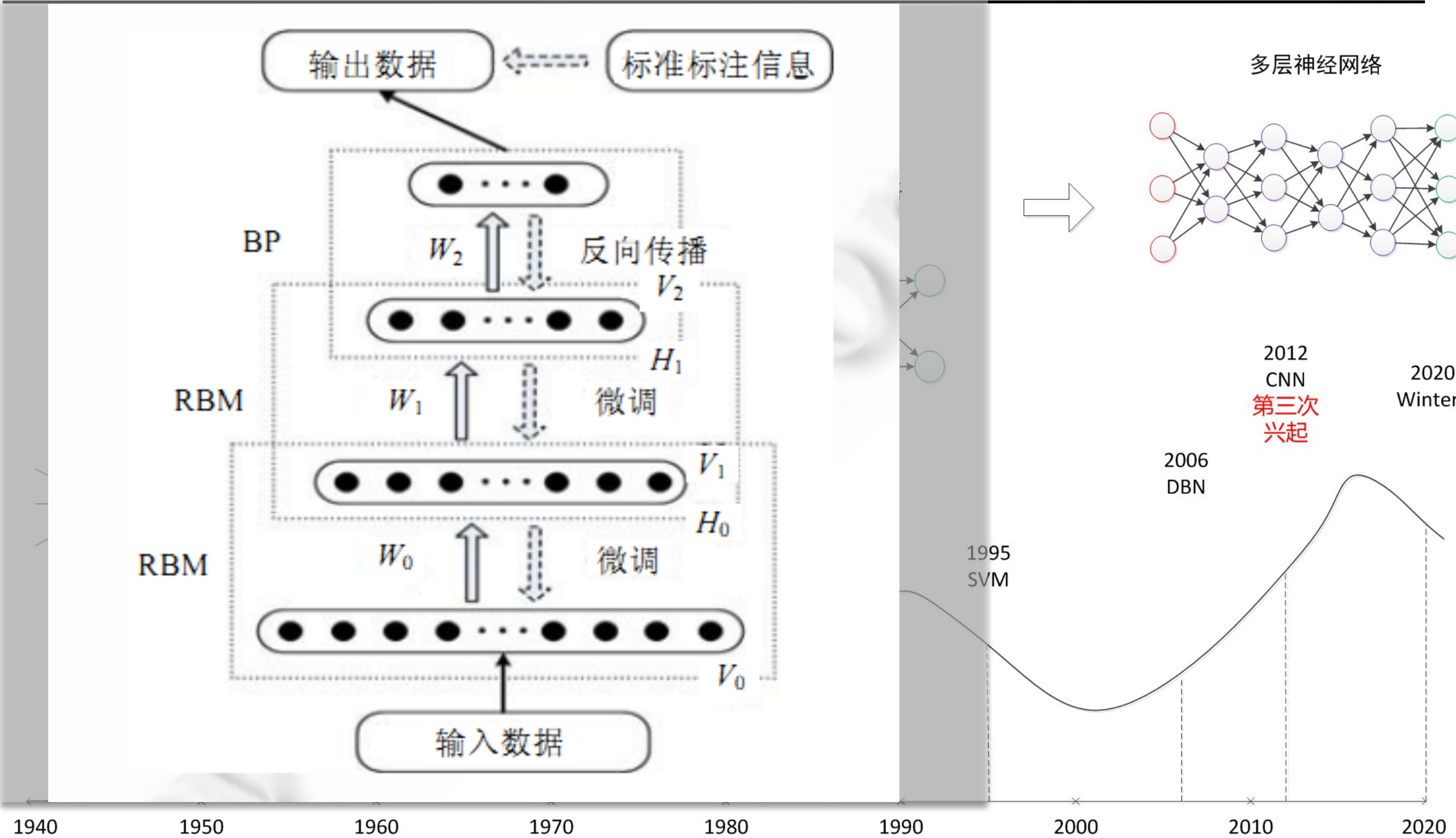
两层神经网络



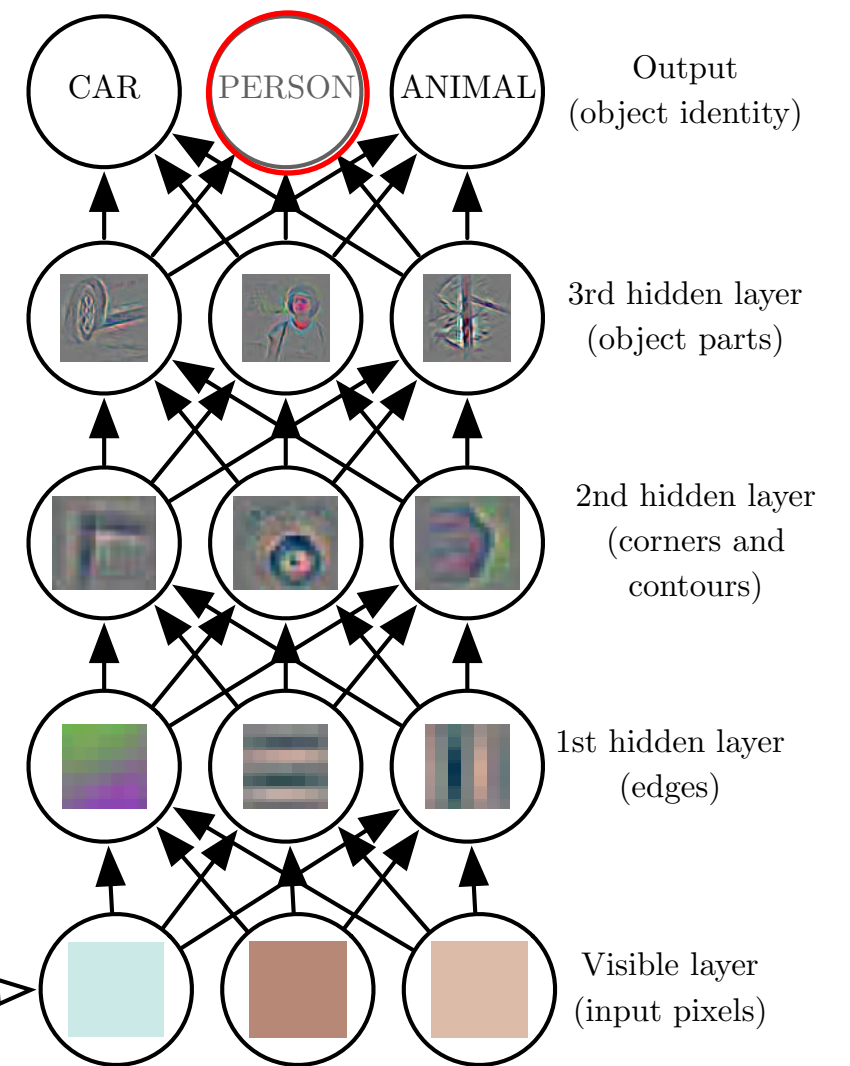
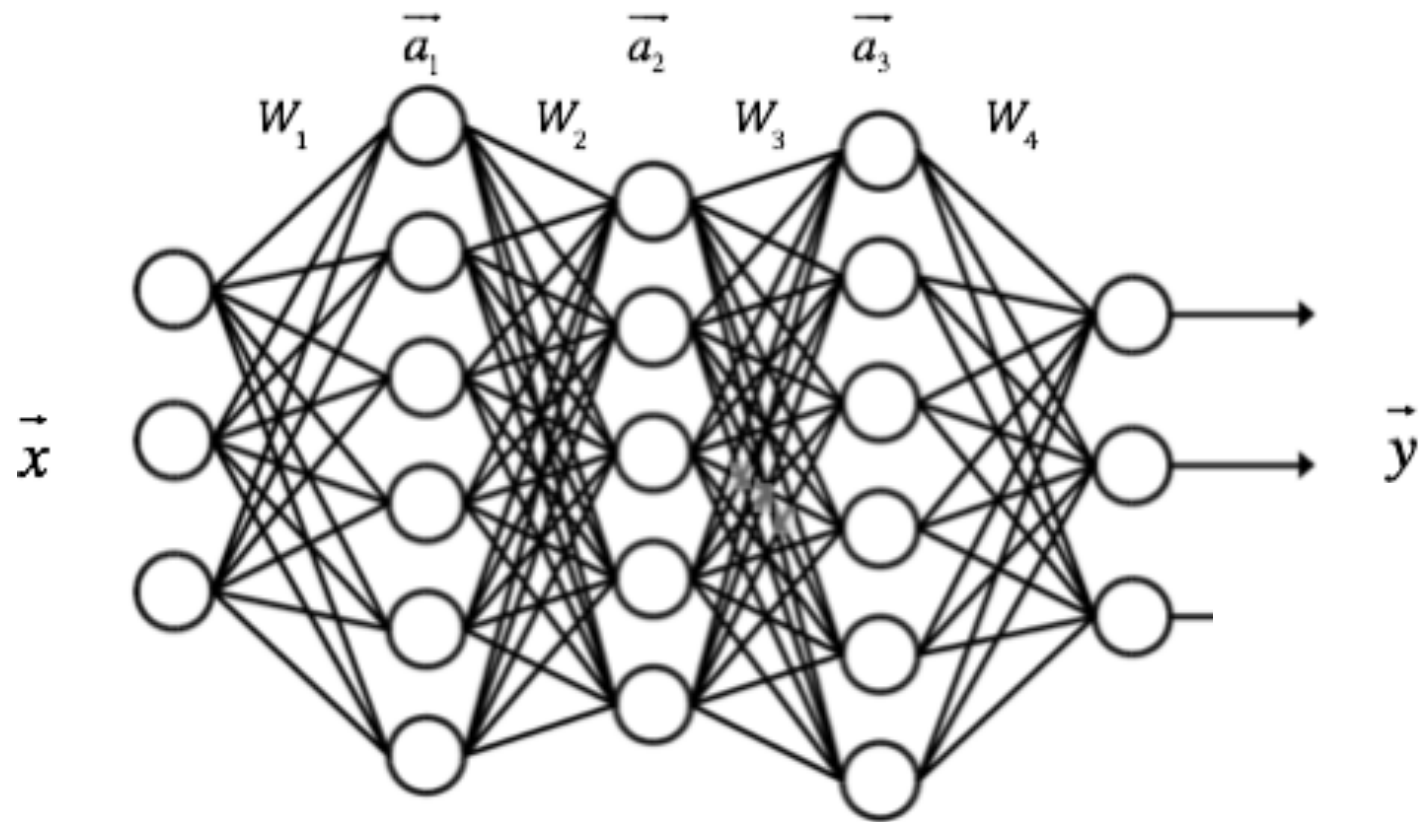
两层神经网络的作用



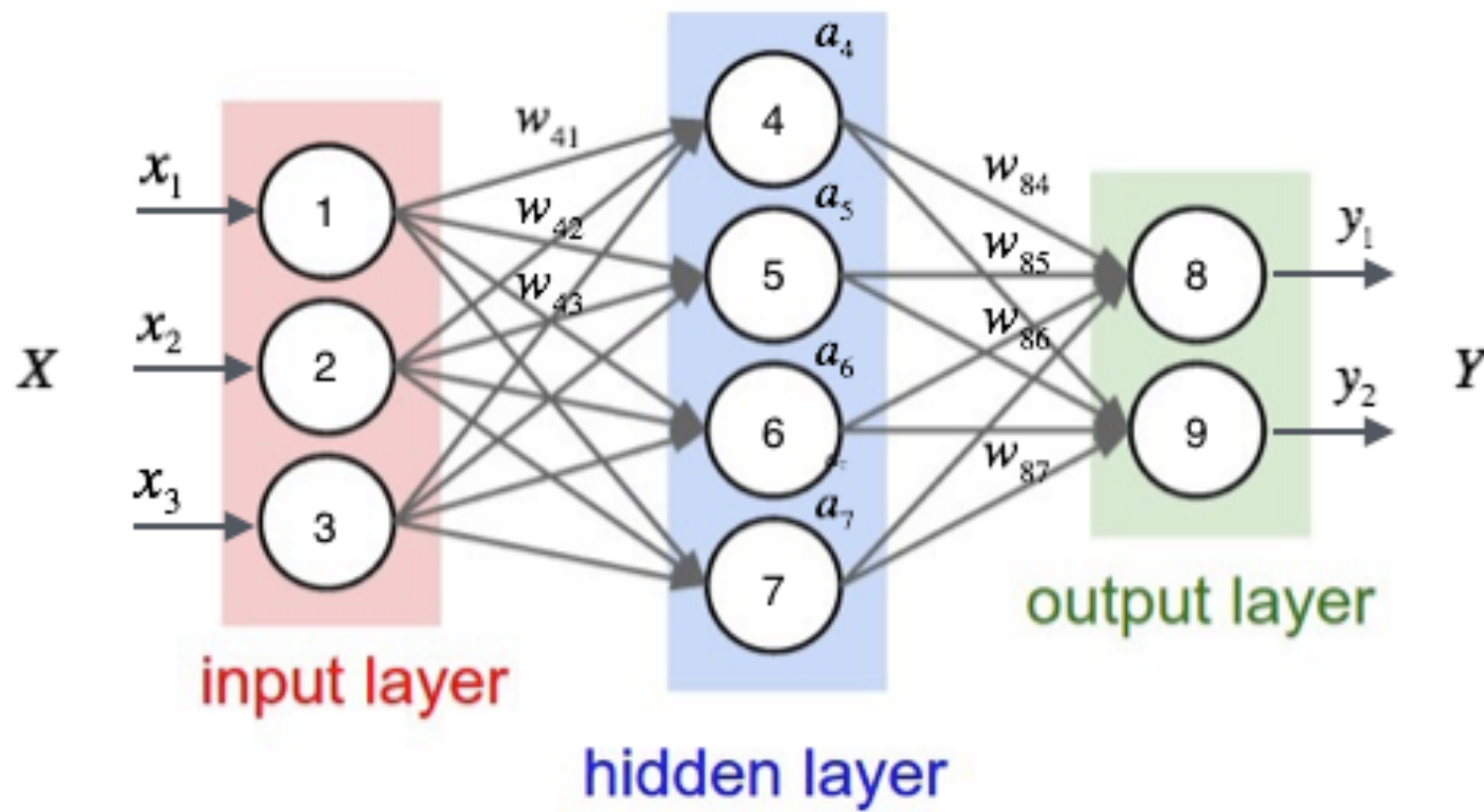
神经网络的历史发展



多层神经网络



神经网络的计算



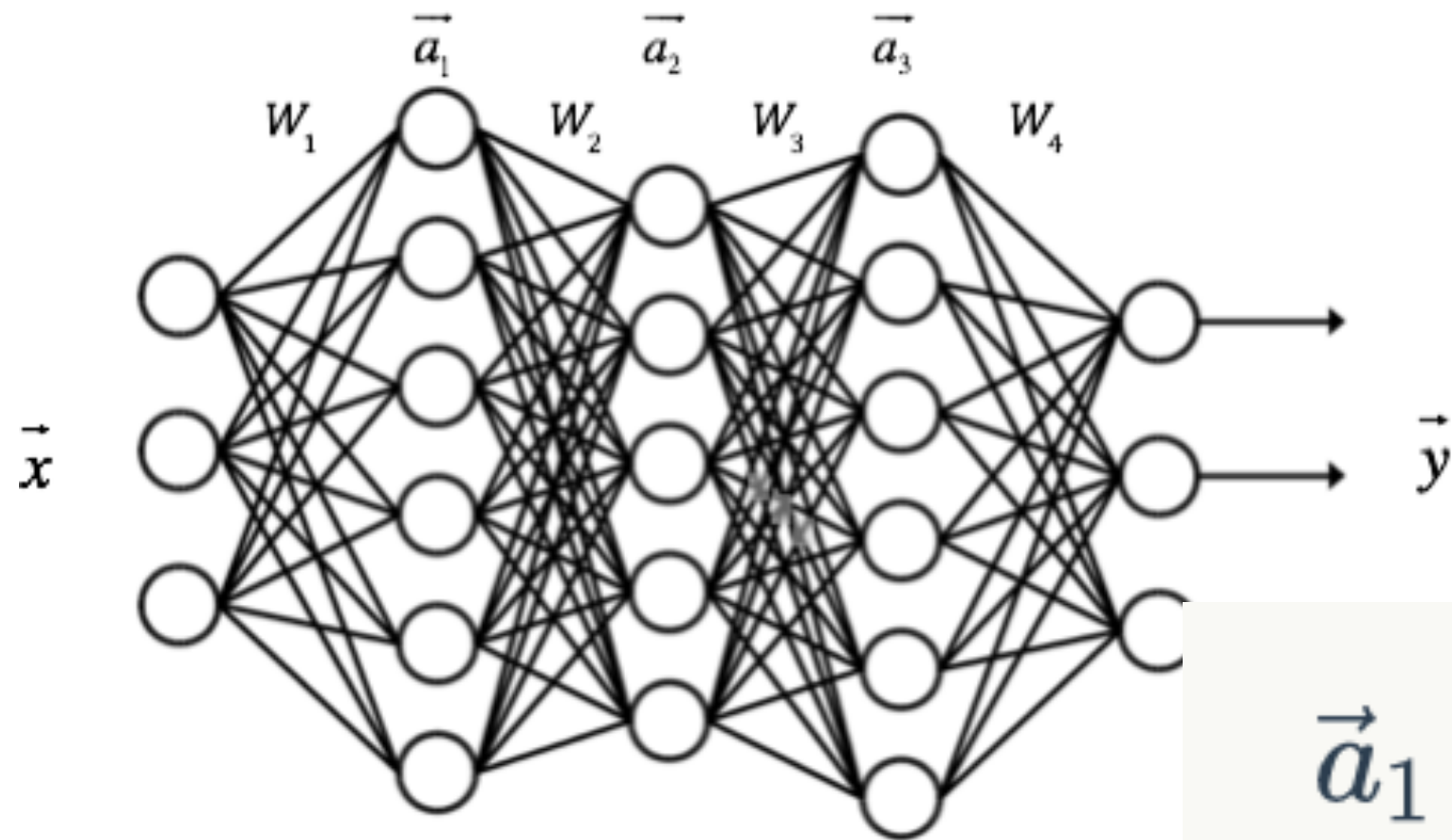
$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix}$$
$$\vec{w}_4 = [w_{41}, w_{42}, w_{43}, w_{b4}]$$
$$\vec{w}_5 = [w_{51}, w_{52}, w_{53}, w_{b5}]$$
$$\vec{w}_6 = [w_{61}, w_{62}, w_{63}, w_{b6}]$$
$$\vec{w}_7 = [w_{71}, w_{72}, w_{73}, w_{b7}]$$
$$f = \text{sigmoid}$$

$$a_4 = \text{sigmoid}(w_{41}x_1 + w_{42}x_2 + w_{43}x_3 + w_{4b})$$
$$a_5 = \text{sigmoid}(w_{51}x_1 + w_{52}x_2 + w_{53}x_3 + w_{5b})$$
$$a_6 = \text{sigmoid}(w_{61}x_1 + w_{62}x_2 + w_{63}x_3 + w_{6b})$$
$$a_7 = \text{sigmoid}(w_{71}x_1 + w_{72}x_2 + w_{73}x_3 + w_{7b})$$



$$a_4 = f(\vec{w}_4 \cdot \vec{x})$$
$$a_5 = f(\vec{w}_5 \cdot \vec{x})$$
$$a_6 = f(\vec{w}_6 \cdot \vec{x})$$
$$a_7 = f(\vec{w}_7 \cdot \vec{x})$$

神经网络的计算



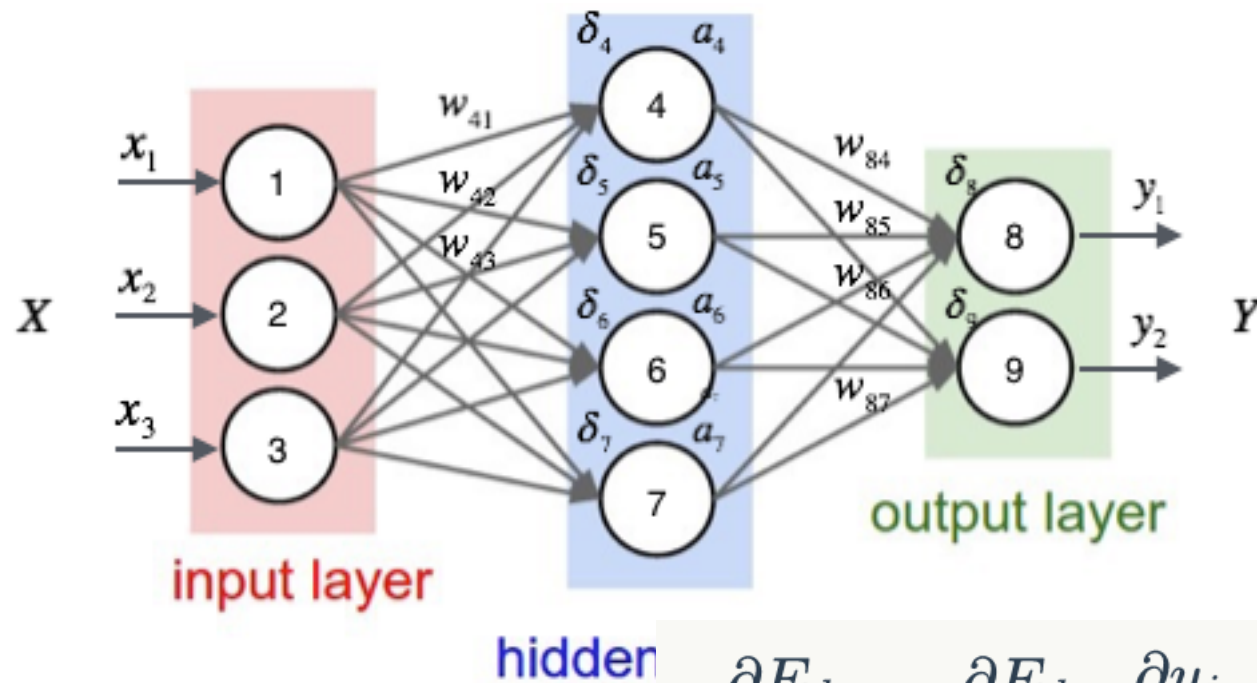
$$\vec{a}_1 = f(W_1 \cdot \vec{x})$$

$$\vec{a}_2 = f(W_2 \cdot \vec{a}_1)$$

$$\vec{a}_3 = f(W_3 \cdot \vec{a}_2)$$

$$\vec{y} = f(W_4 \cdot \vec{a}_3)$$

反向传播算法



$$E_d \equiv \frac{1}{2} \sum_{i \in \text{outputs}} (t_i - y_i)^2$$

$$w_{ji} \leftarrow w_{ji} - \eta \frac{\partial E_d}{\partial w_{ji}}$$

$$\begin{aligned} net_j &= \vec{w}_j \cdot \vec{x}_j \\ &= \sum_i w_{ji} x_{ji} \end{aligned}$$

$$\begin{aligned} \frac{\partial E_d}{\partial net_j} &= \frac{\partial E_d}{\partial y_j} \frac{\partial y_j}{\partial net_j} \\ &= \frac{\partial E_d}{\partial y_j} \frac{\partial y_j}{\partial net_j} \end{aligned}$$

$$\frac{\partial E_d}{\partial net_j} = -(t_j - y_j) y_j (1 - y_j)$$

$$\delta_i = y_i (1 - y_i) (t_i - y_i)$$

$$\begin{aligned} \frac{\partial E_d}{\partial w_{ji}} &= \frac{\partial E_d}{\partial net_j} \frac{\partial net_j}{\partial w_{ji}} \\ &= \frac{\partial E_d}{\partial net_j} \frac{\partial \sum_i w_{ji} x_{ji}}{\partial w_{ji}} \\ &= \frac{\partial E_d}{\partial net_j} x_{ji} \end{aligned}$$

$$\delta_i = a_i (1 - a_i) \sum_{k \in \text{outputs}} w_{ki} \delta_k + \eta \delta_j x_{ji}$$

THANK YOU

谢谢聆听

ThoughtWorks®