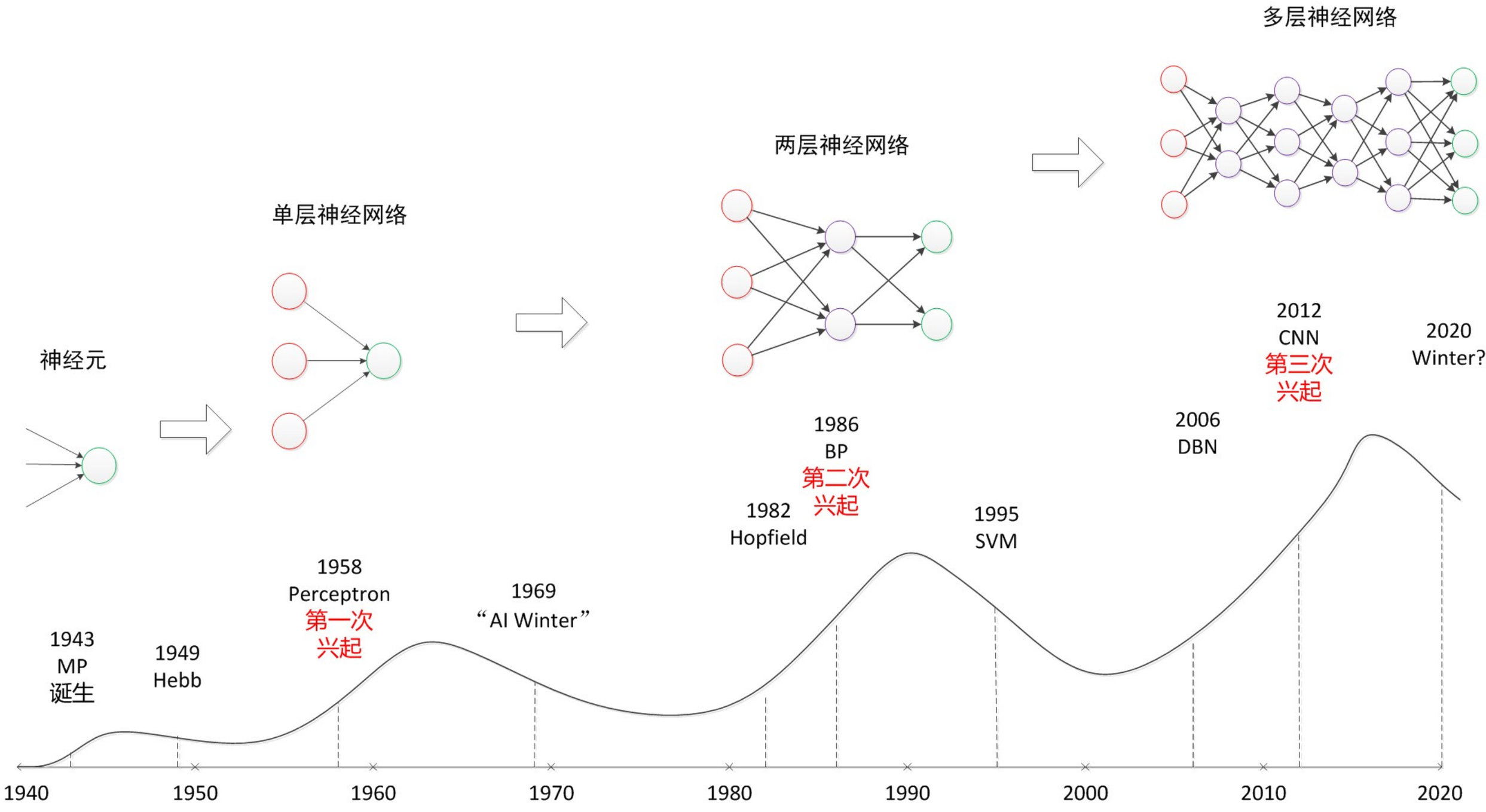
ThoughtWorks®

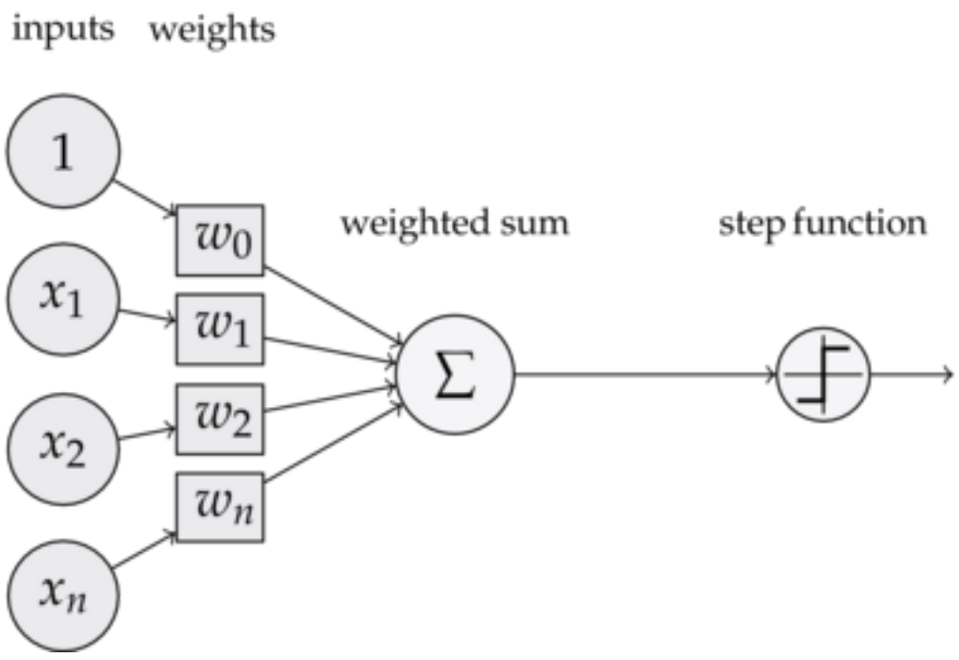
Neural Network

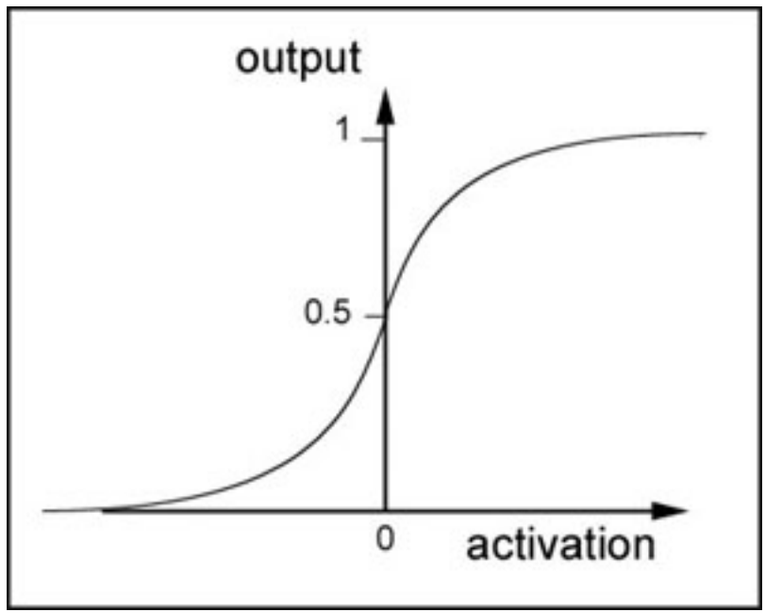
INTRODUCTION

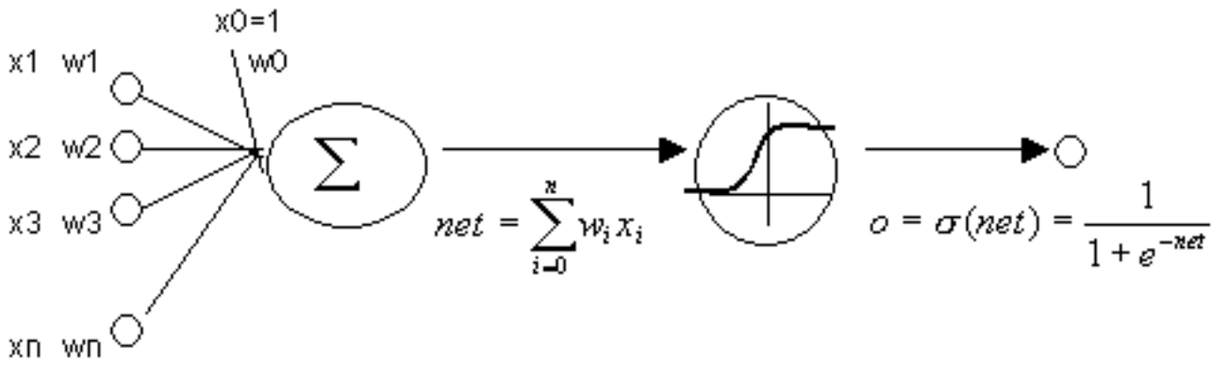
THE HISTORY OF NEURAL NETWORK



WHAT'S THE NEURON





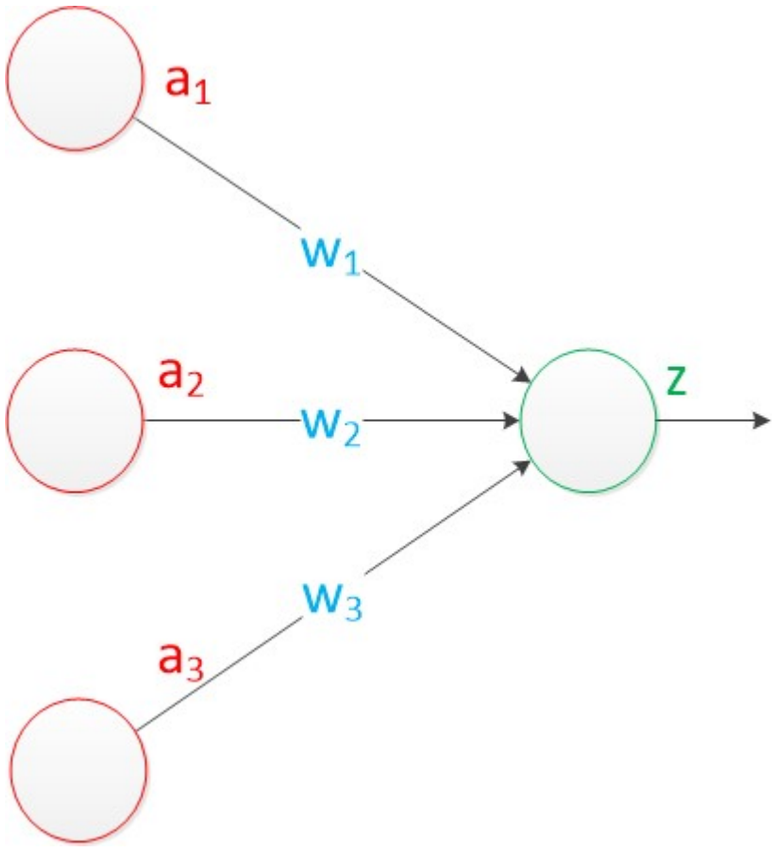


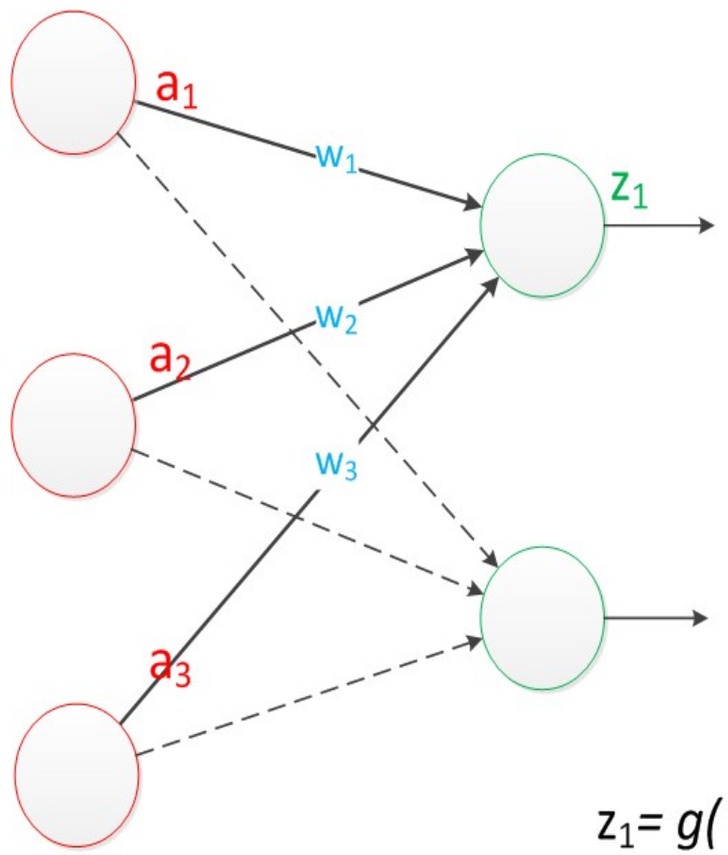
[The sigmoid threshold unit]

令$y = sigmoid(x)$

则$y' = y(1-y)$

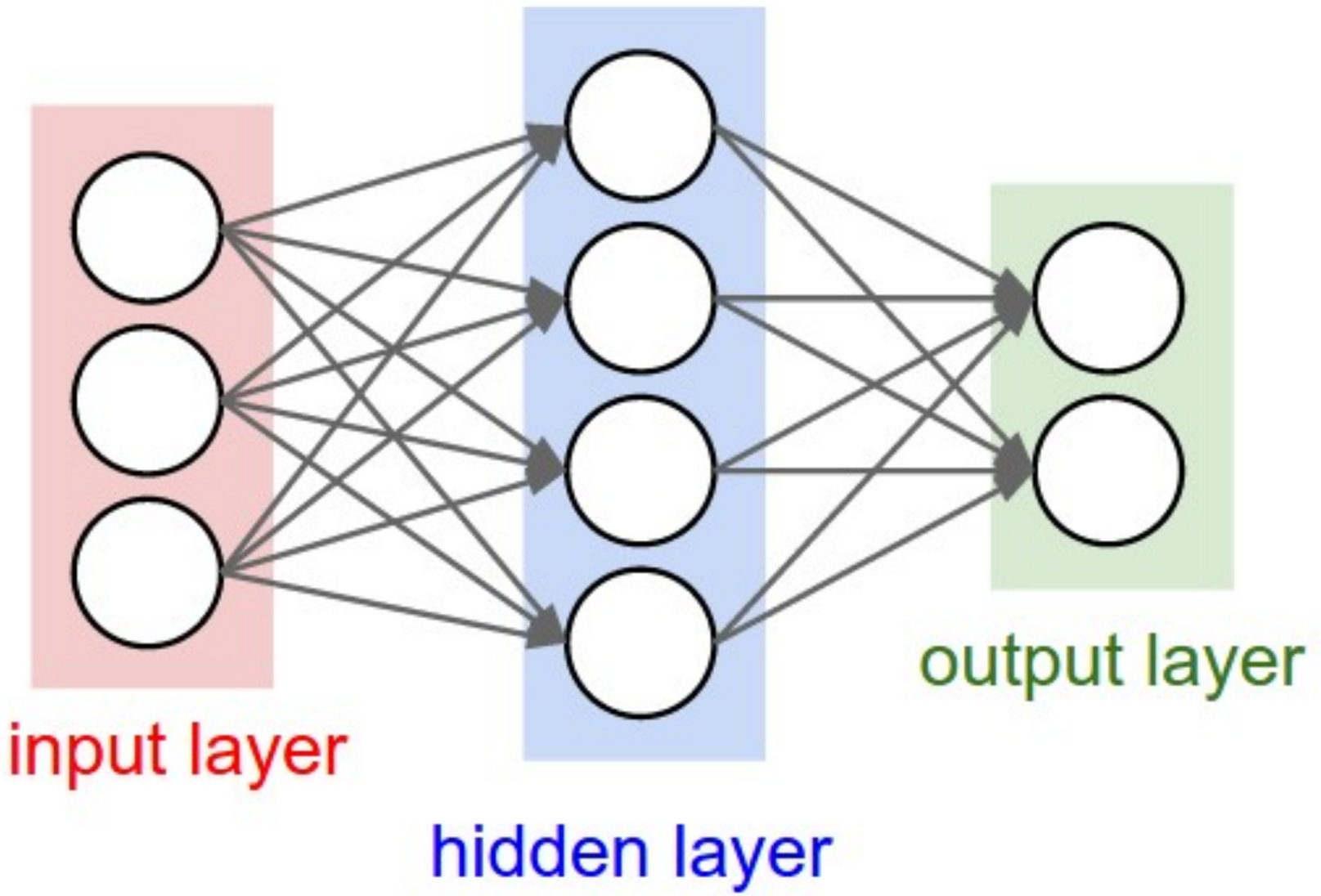
WHAT'S THE 1-NEURAL NETWORK



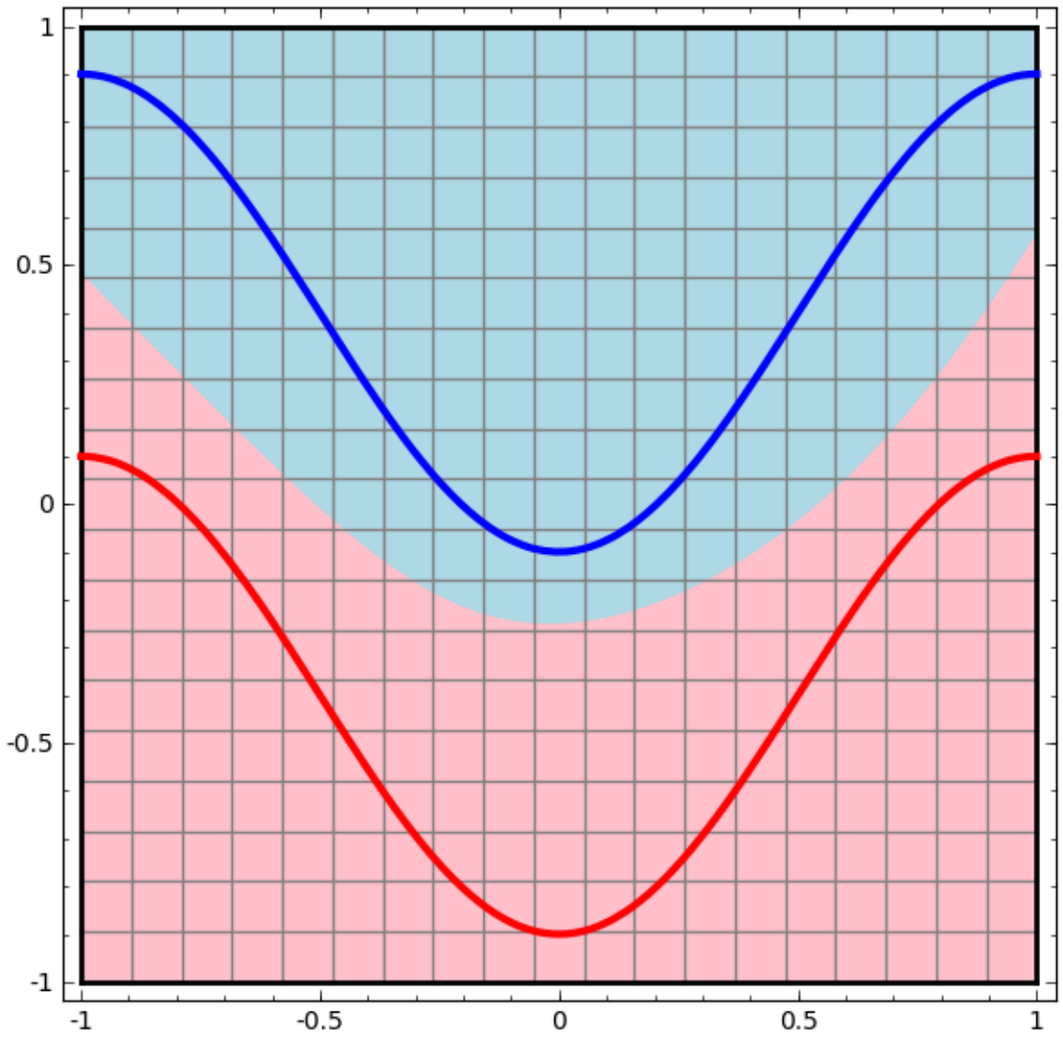


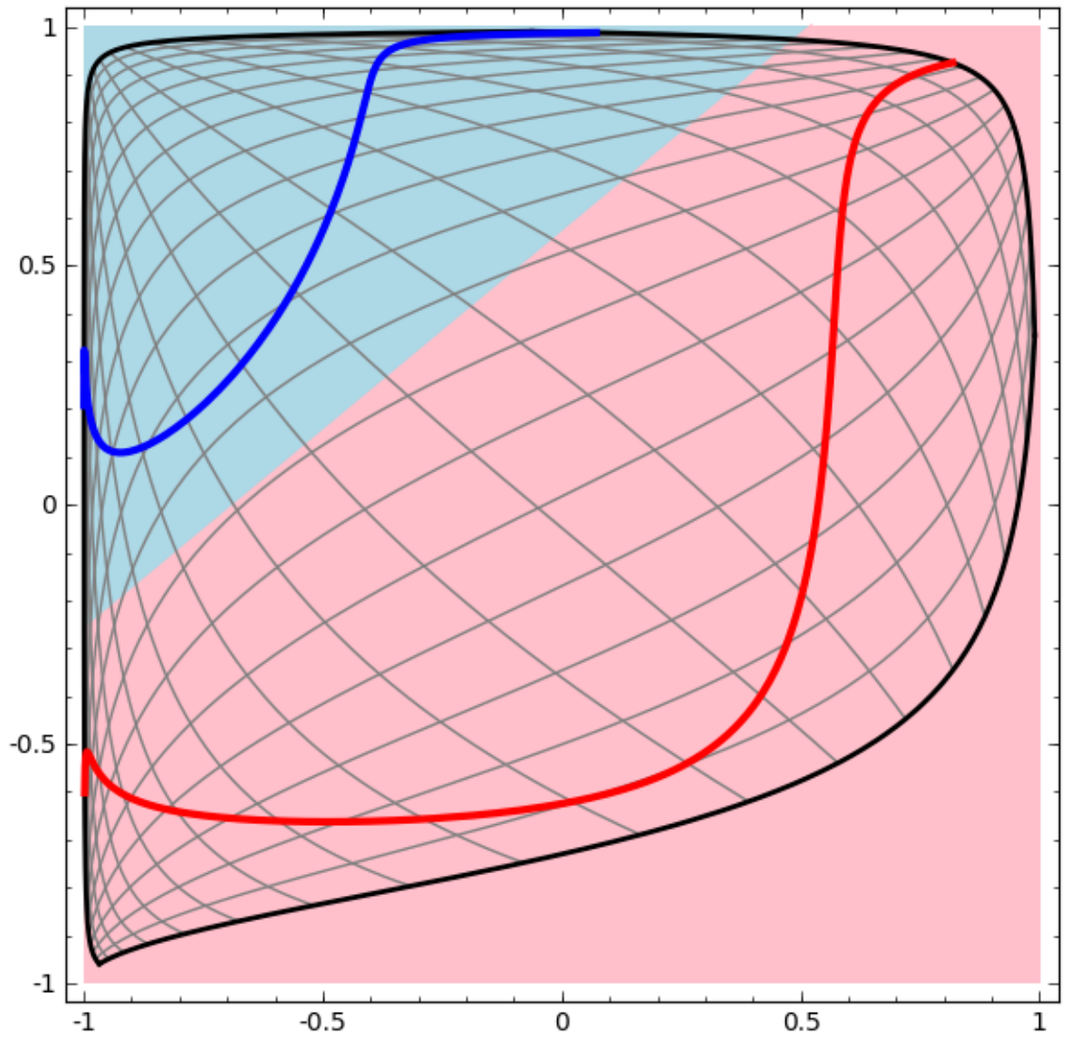
$z_1 = g(a_1 * w_1 + a_2 * w_2 + a_3 * w_3)$

WHAT'S THE 2-NEURAL NETWORK

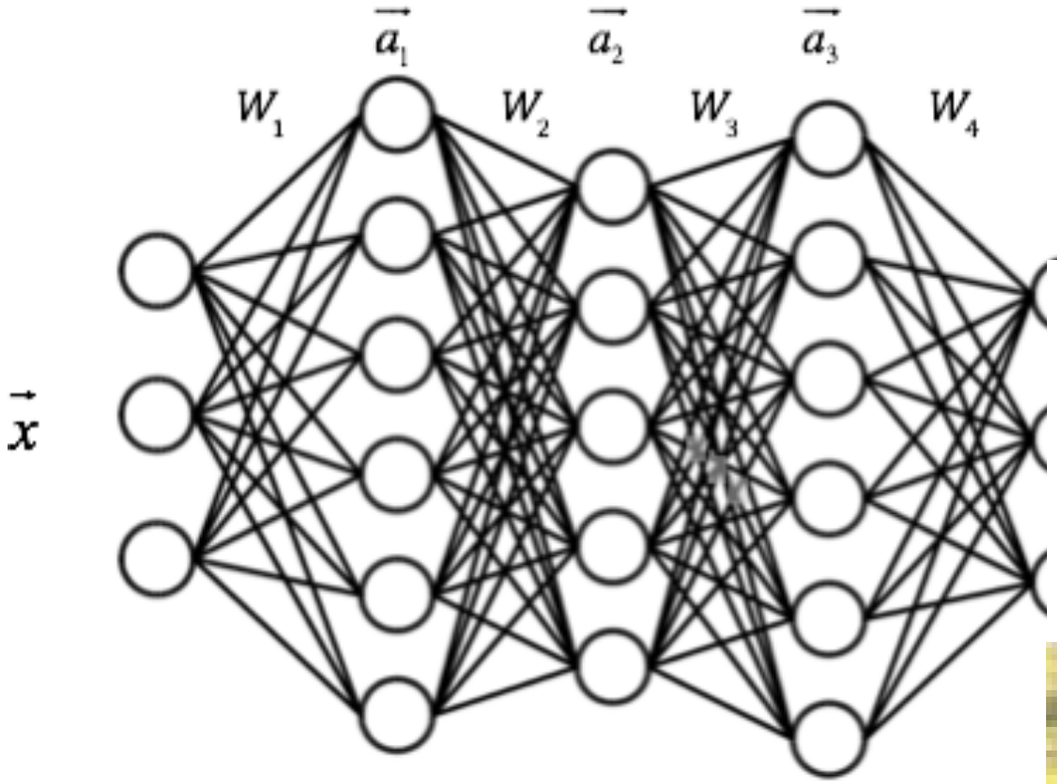


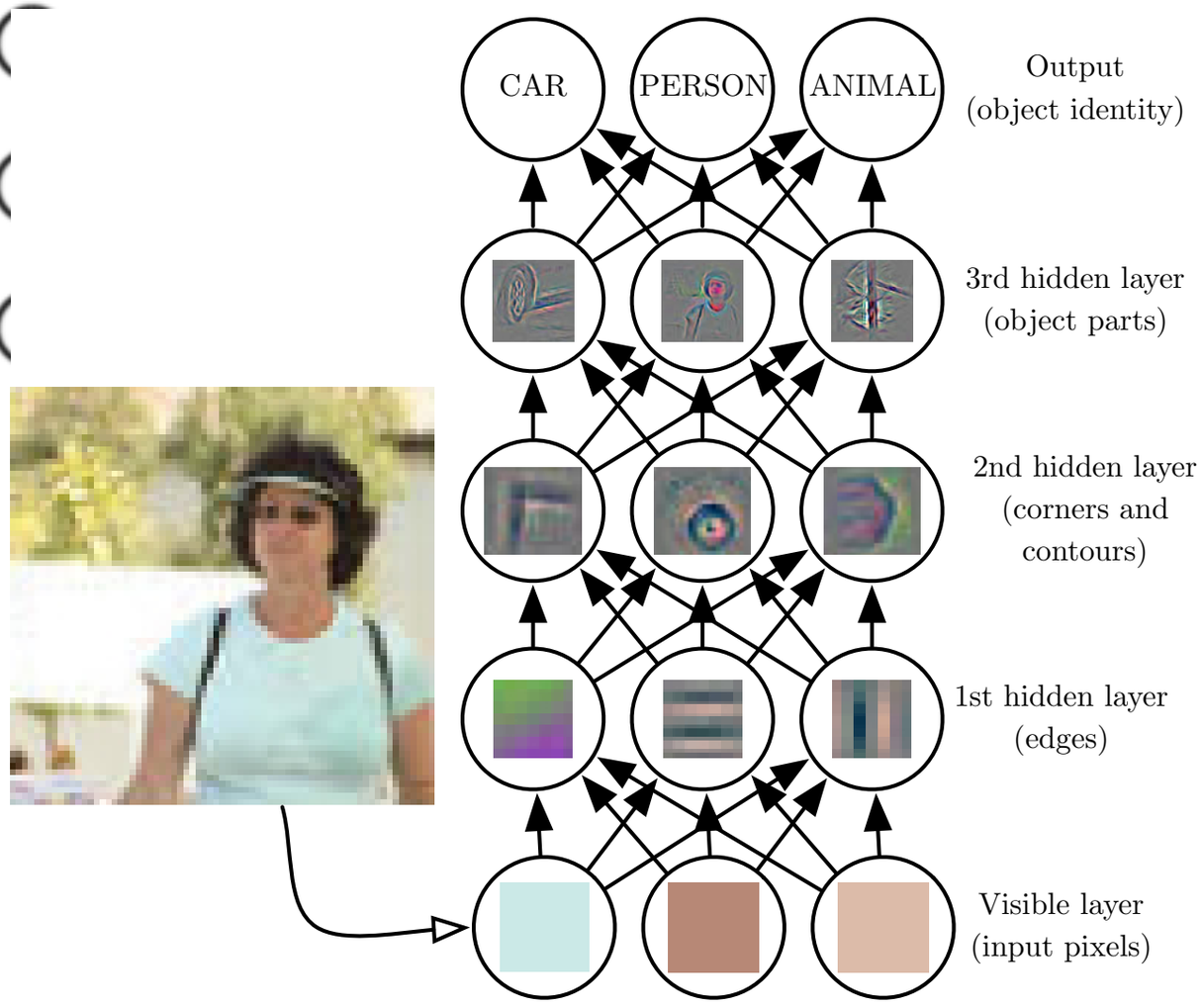
WHAT'S THE EFFECT 2-NEURAL NETWORK



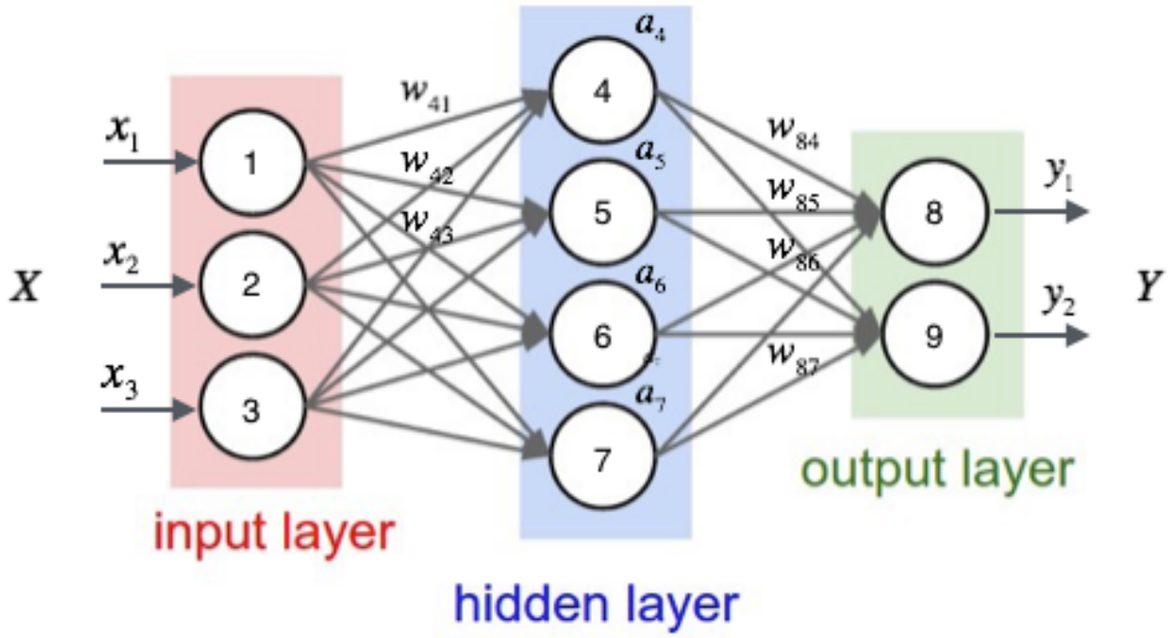


WHAT'S THE EFFECT MULTI-NEURAL NETWORK





WHAT'S THE OUTPUT OF NEURAL NETWORK



$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix}$$

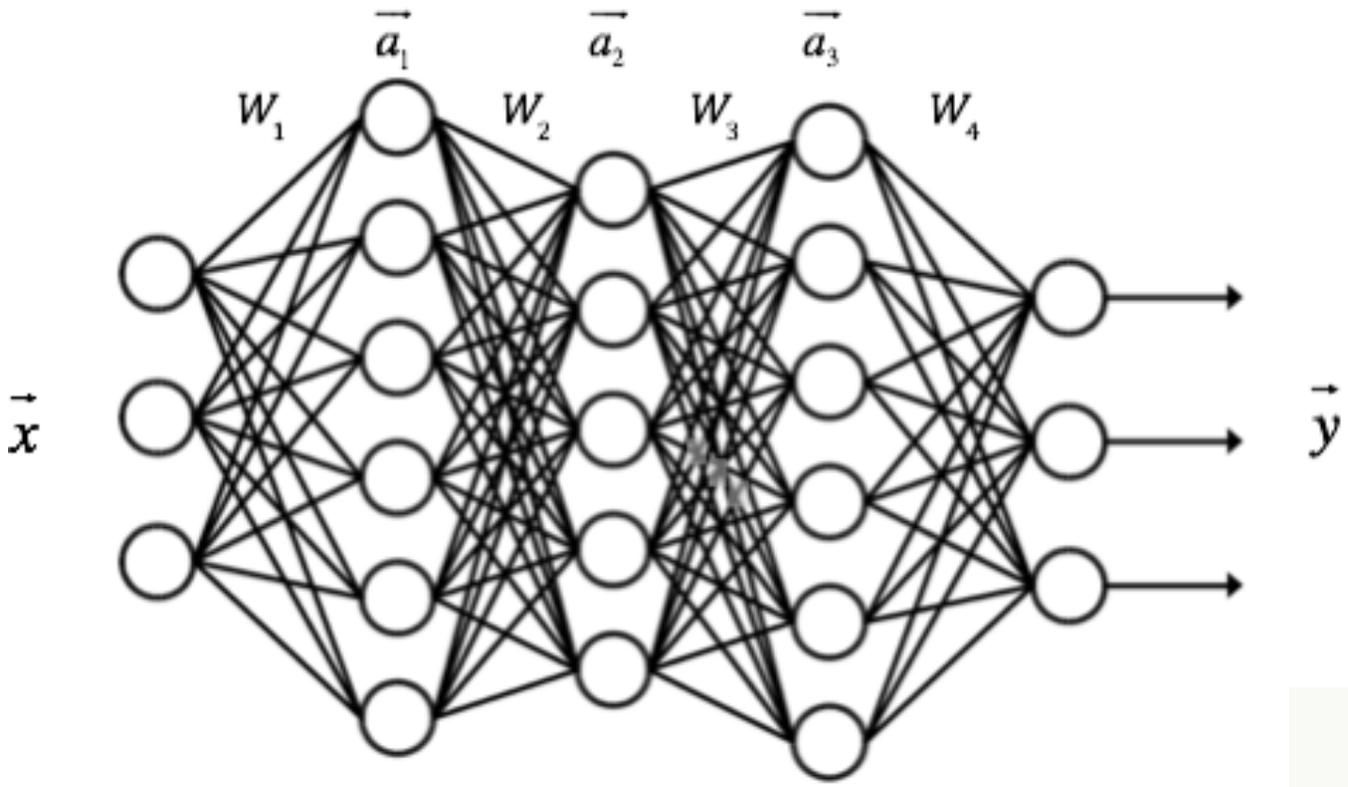
$$\vec{w}_4 = [w_{41}, w_{42}, w_{43}, w_{b4}]$$
$$\vec{w}_5 = [w_{51}, w_{52}, w_{53}, w_{b5}]$$
$$\vec{w}_6 = [w_{61}, w_{62}, w_{63}, w_{b6}]$$
$$\vec{w}_7 = [w_{71}, w_{72}, w_{73}, w_{b7}]$$

$$f = sigmoid$$

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

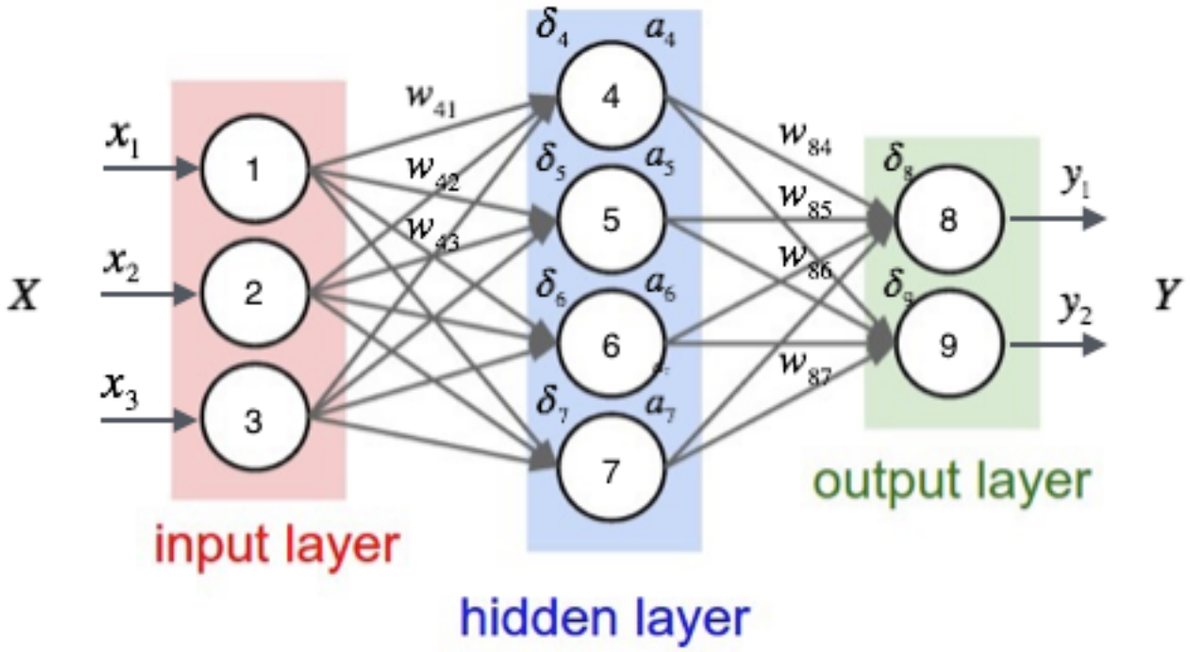
$$a_4 = f(\vec{w_4} \cdot \vec{x})$$
$$a_5 = f(\vec{w_5} \cdot \vec{x})$$
$$a_6 = f(\vec{w_6} \cdot \vec{x})$$
$$a_7 = f(\vec{w_7} \cdot \vec{x})$$

WHAT'S THE OUTPUT OF NEURAL NETWORK



$$\vec{a}_1 = f(W_1 \cdot \vec{x})$$
$$\vec{a}_2 = f(W_2 \cdot \vec{a}_1)$$
$$\vec{a}_3 = f(W_3 \cdot \vec{a}_2)$$
$$\vec{y} = f(W_4 \cdot \vec{a}_3)$$

WHAT'S THE BACK PROPAGATION



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

$$\delta_8 = y_1(1-y_1)(t_1 - y_1)$$

$$\delta_i = a_i(1-a_i)\sum_{k \in outputs} w_{ki}\delta_k$$

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

$$\delta_4 = a_4(1-a_4)(w_{84}\delta_8 + w_{94}\delta_9)$$

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

$$w_{ji} \leftarrow w_{ji} + \eta\delta_j x_{ji}$$

$$net_j = \vec{w_j} \cdot \vec{x_j} = \sum_i w_{ji}x_{ji}$$

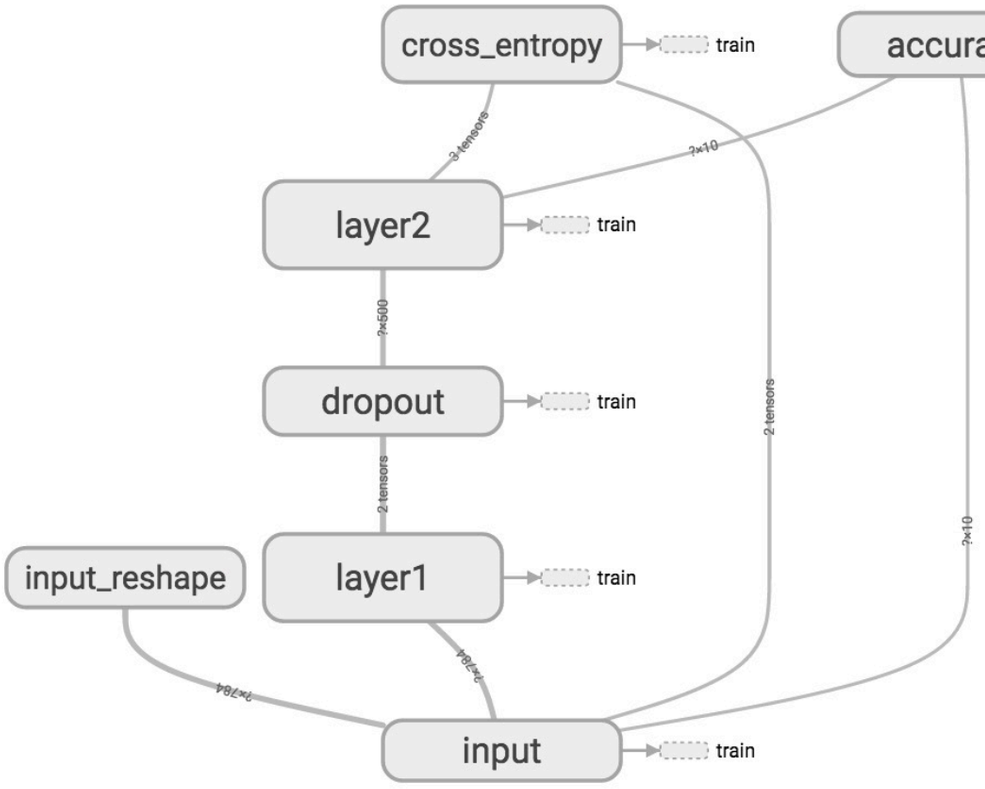
$$\frac{\partial E_d}{\partial w_{ji}} = \frac{\partial E_d}{\partial net_j}\frac{\partial net_j}{\partial w_{ji}} = \frac{\partial E_d}{\partial net_j}\frac{\partial \sum_i w_{ji}x_{ji}}{\partial w_{ji}} = \frac{\partial E_d}{\partial net_j}x_{ji}$$

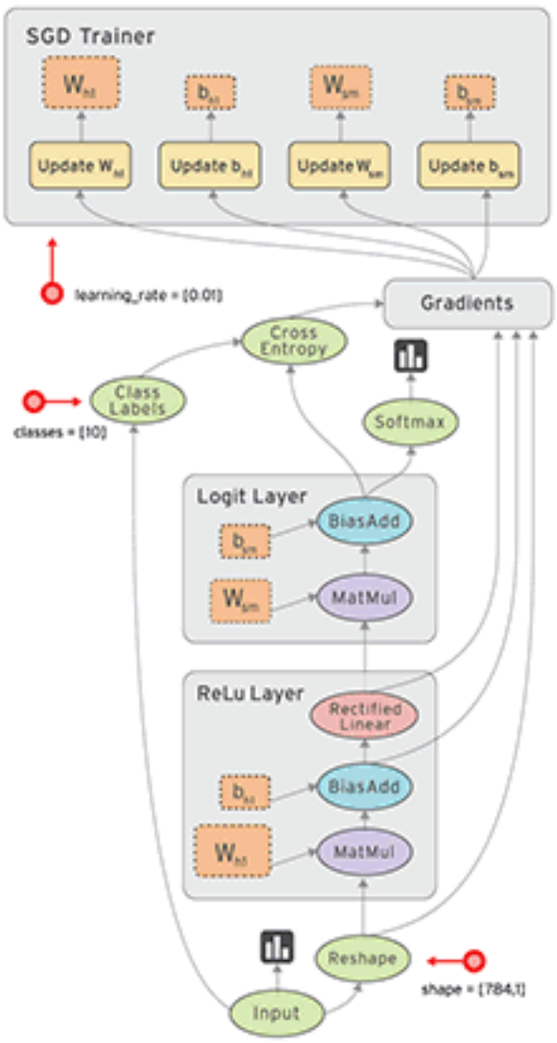
$$\frac{\partial E_d}{\partial net_j} = \frac{\partial E_d}{\partial y_j}\frac{\partial y_j}{\partial net_j} = \frac{\partial E_d}{\partial y_j}\frac{\partial y_j}{\partial net_j}$$

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

WHAT'S ABOUT TENSORFLOW

- Represents computations as graphs.
- Executes graphs in the context of Sessions.
- Represents data as tensors.
- Maintains state with Variables.
- Uses feeds and fetches to get data into and out of arbitrary operations.





THANK YOU

Please give me feedbacks

ThoughtWorks®