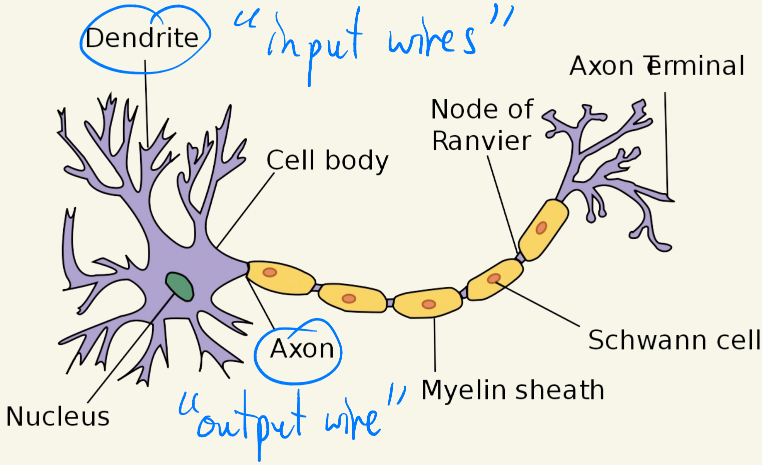
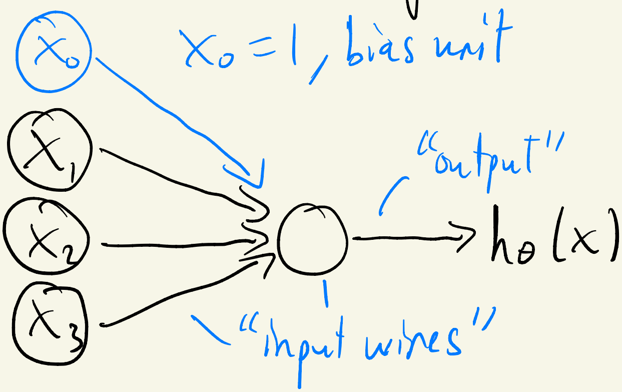
Nerve Cell (Neuron):

May 18, 2021



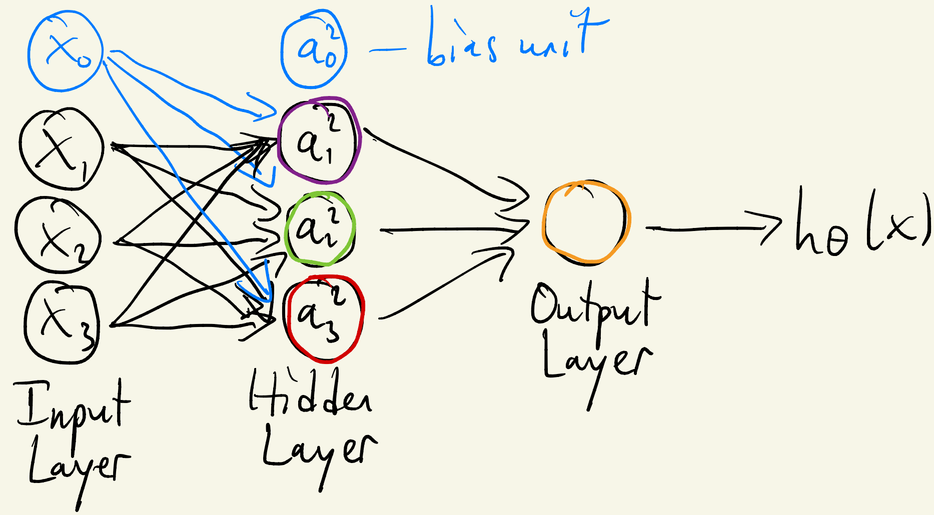
Neurons send signals via electrical signals from their output wires to another neuron's input wires.

Neuron Model (Logistic Unit):



$$h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$$

Neural Network:



$a_i^{(j)}$ – "activation" of unit i in layer j

$\Theta^{(j)}$ – matrix of weights controlling function mapping from layer j to $j+1$

i.e.

$$a_1^{(2)} = g(\underbrace{\Theta_{10}^{(1)} x_0 + \Theta_{11}^{(1)} x_1 + \Theta_{12}^{(1)} x_2 + \Theta_{13}^{(1)} x_3}_{z_1^{(2)}})$$

$$a_2^{(2)} = g(\underbrace{\Theta_{20}^{(1)} x_0 + \Theta_{21}^{(1)} x_1 + \Theta_{22}^{(1)} x_2 + \Theta_{23}^{(1)} x_3}_{z_2^{(2)}})$$

$$a_3^{(2)} = g(\underbrace{\Theta_{30}^{(1)} x_0 + \Theta_{31}^{(1)} x_1 + \Theta_{32}^{(1)} x_2 + \Theta_{33}^{(1)} x_3}_{z_3^{(2)}})$$

$$h_\theta(x) = a_1^{(3)} = g(\Theta_{10}^{(2)} a_0^{(2)} + \Theta_{11}^{(2)} a_1^{(2)} + \Theta_{12}^{(2)} a_2^{(2)} + \Theta_{13} a_3^{(2)})$$

Shape Θ^j = (# hidden layer neurons) $\times$ (# inputs + 1) i.e.,

$$\Theta^{(1)} = \begin{bmatrix} \Theta_{10} & \Theta_{11} & \Theta_{12} & \Theta_{13} \\ \Theta_{20} & \Theta_{21} & \Theta_{22} & \Theta_{23} \\ \Theta_{30} & \Theta_{31} & \Theta_{32} & \Theta_{33} \end{bmatrix}$$

shape $= s_{j+1} \times (s_j + 1)$
where $s = \#$ of units

3×4

$$x = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad z^{(2)} = \begin{bmatrix} z_1^{(2)} \\ z_2^{(2)} \\ z_3^{(2)} \end{bmatrix}$$

$z^{(2)} = \Theta^{(1)} x$
$a^{(2)} = g(z^{(2)})$

Add $a_0^{(2)} = 1$.
$z^{(3)} = \Theta^{(2)} a^{(2)}$
$h_\theta(x) = a^{(3)} = g(z^{(3)})$

* This is known as forward propagation

* Compute activations from input, hidden, and output layers.

Notice the Similarity!

$$h_\Theta(x) = g(\Theta_{10}^{(2)} a_0^{(2)} + \Theta_{11}^{(2)} a_1^{(2)} + \Theta_{12}^{(2)} a_2^{(2)} + \Theta_{13}^{(2)} a_3^{(2)})$$

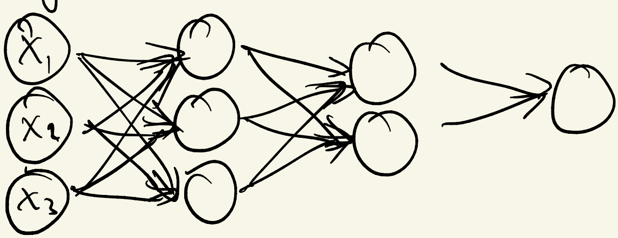
↳ like logistic regression, but uses $a_0 \ldots a_3$ instead of $x_0 \ldots x_3$

* Neural networks are not constrained to feeding $x_0 \ldots x_3$ to logistic regression, but can feed its own trained features $a_0 \ldots a_3$ into logistic regression.

↳ Very flexible, can learn some very complex features

The neural network's architecture refers to how the neurons connect to each other

e.g.



Different architecture from before.