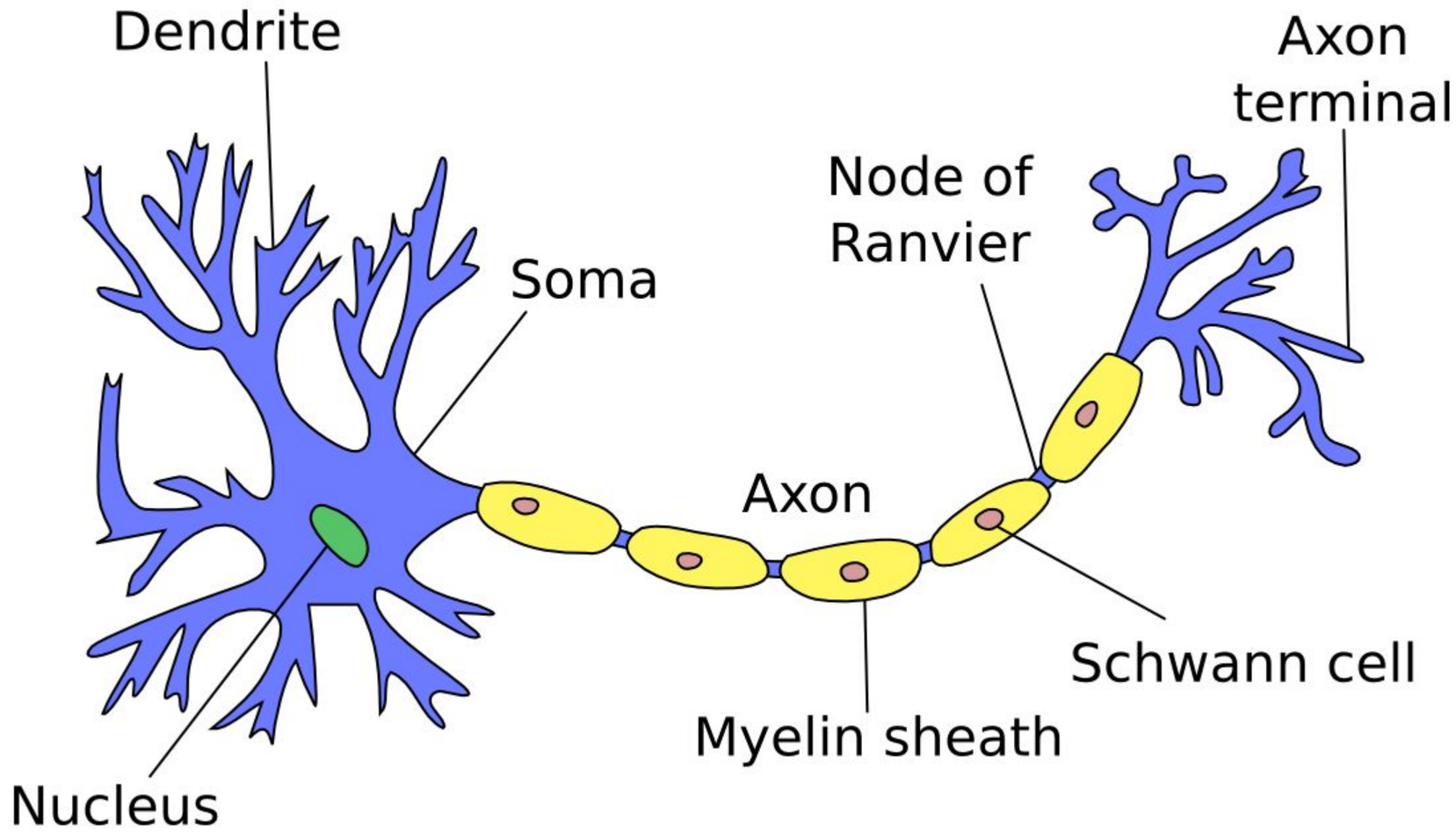
Deep Learning

Big Data & Machine Learning Bootcamp - Keep Coding

Outline

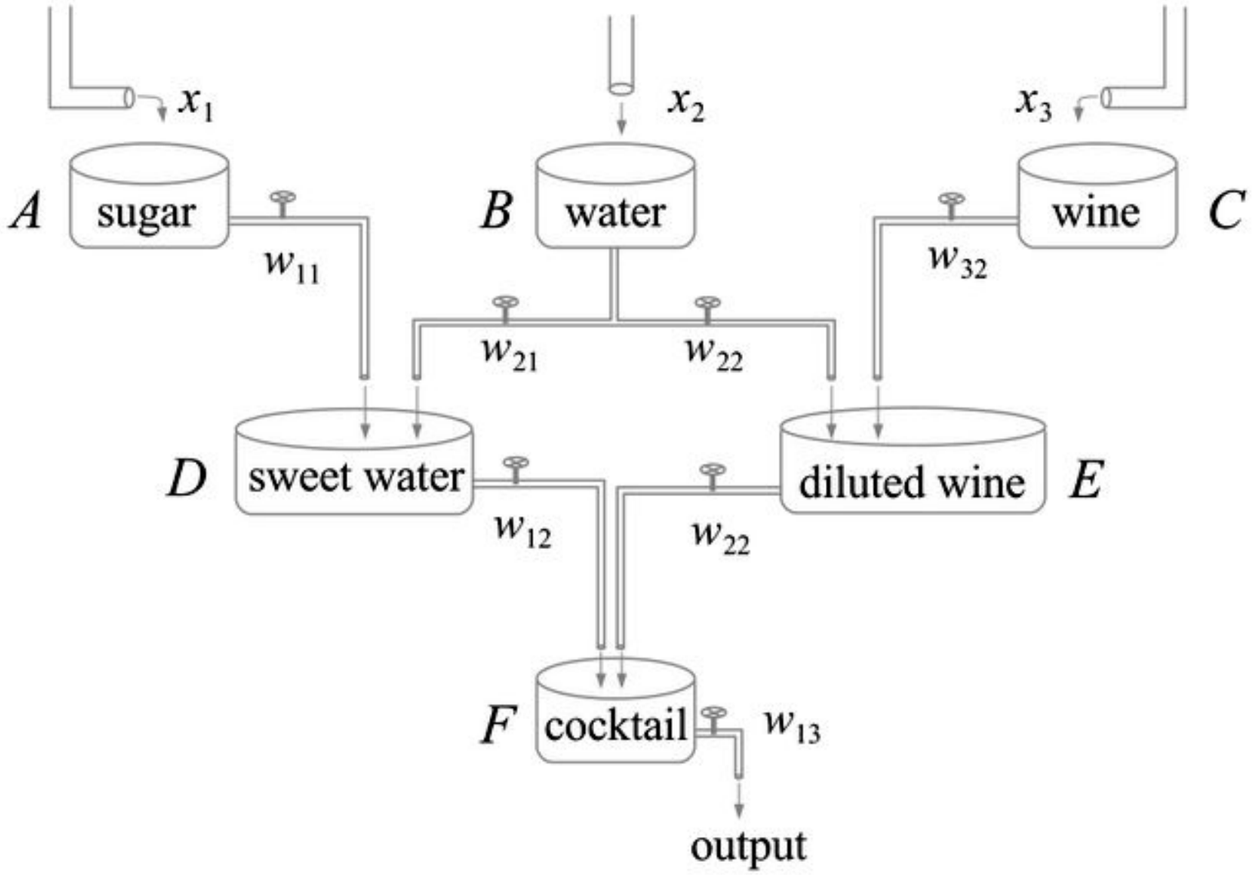
1. Neuron and its linear representation
2. Hidden layer
3. Activation functions
4. Forward and backward propagation

Source of inspiration



Source: https://www.marekrei.com/blog/neural-networks-part-2-the-neuron/

Examples of neural networks



A cocktail factory as a neural network with two hidden layers

Source: Ovidiu Calin. Deep Learning Architectures: A Mathematical Approach

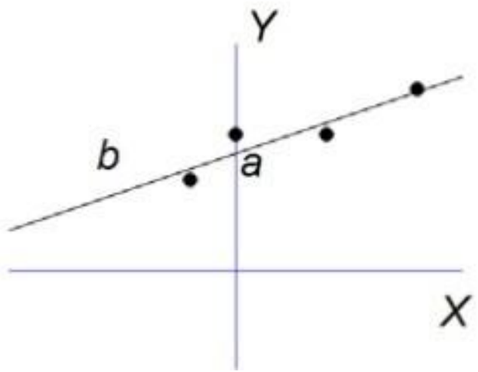
$$\hat{Y} = bX + a$$

predicted values of *Y*

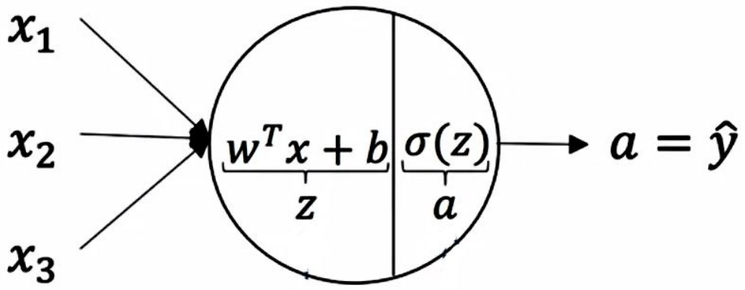
b = slope = rate of predicted ↑/↓ for Y scores for each unit increase in *X*

Y-intercept = level of *Y* when X is 0

Also called bias (b)

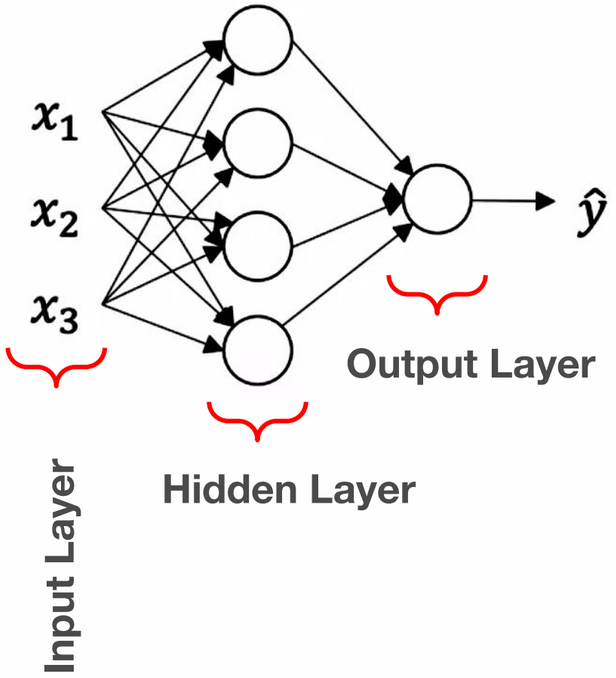


Sources:
- https://oldwomanatschool.blog/2017/10/14/age-is-just-a-perception/

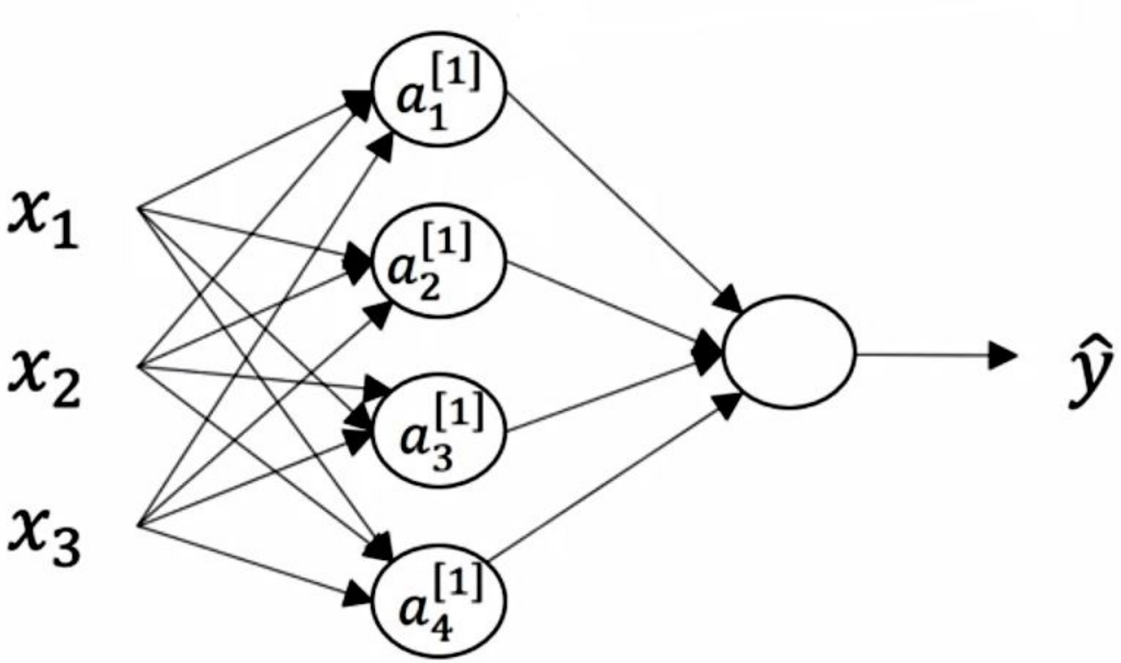


$$z = w^T x + b$$

$$a = \sigma(z)$$



Sources:
- Coursera



$$z_1^{[1]} = w_1^{[1]T} x + b_1^{[1]},\ a_1^{[1]} = \sigma(z_1^{[1]})$$

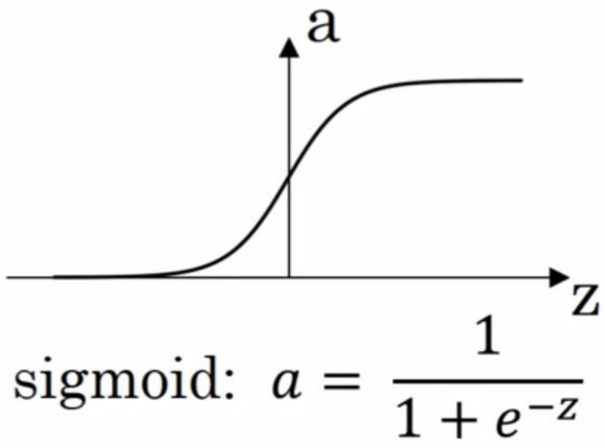
$$z_2^{[1]} = w_2^{[1]T} x + b_2^{[1]},\ a_2^{[1]} = \sigma(z_2^{[1]})$$

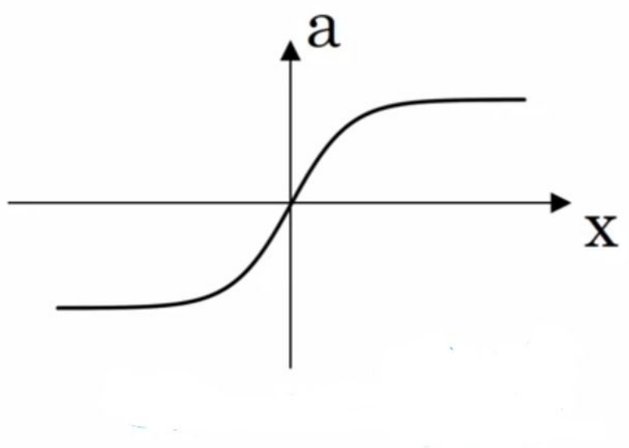
$$z_3^{[1]} = w_3^{[1]T} x + b_3^{[1]},\ a_3^{[1]} = \sigma(z_3^{[1]})$$

$$z_4^{[1]} = w_4^{[1]T} x + b_4^{[1]},\ a_4^{[1]} = \sigma(z_4^{[1]})$$

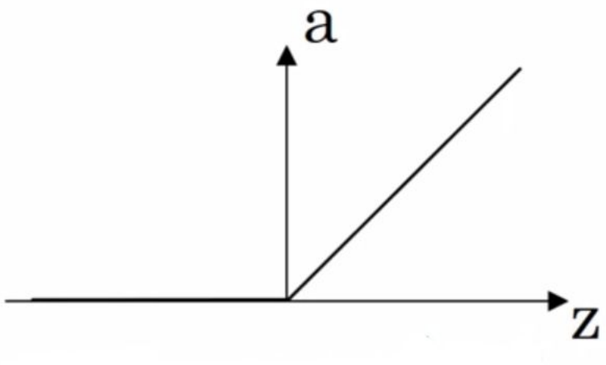
Sources:
- Coursera

Tanh

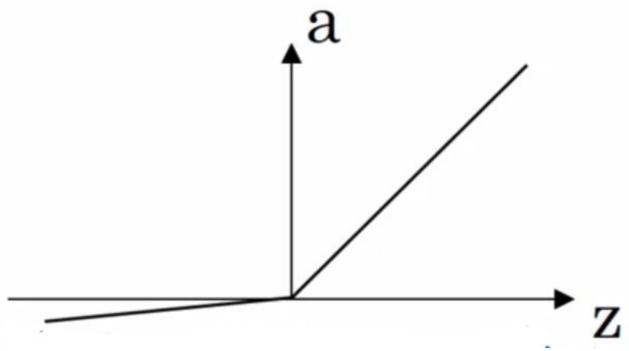




ReLU

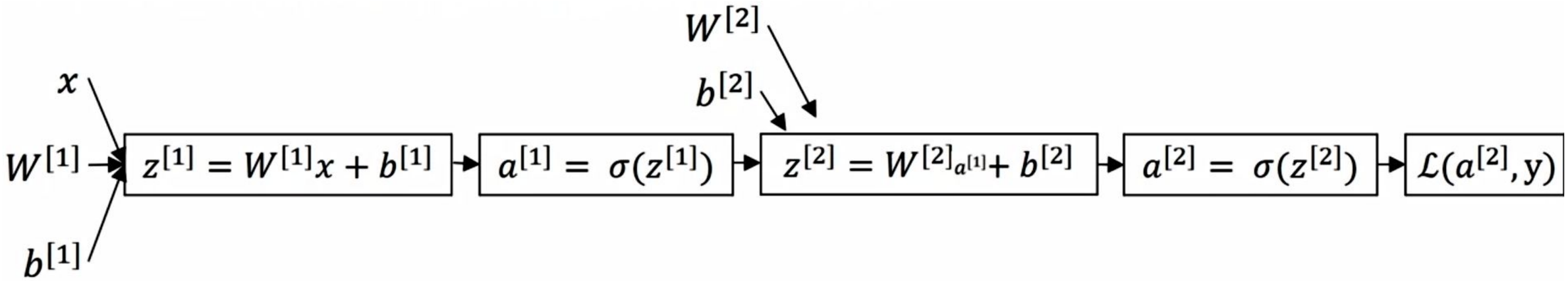


Leaky ReLU



They help for non-linear mapping/functions between the input and the output!

Sources:
- Coursera



Output Layer

$$dz^{[2]} = a^{[2]} - y$$

$$dW^{[2]} = dz^{[2]} a^{[1]^T}$$

$$db^{[2]} = dz^{[2]}$$

Hidden Layer

$$dz^{[1]} = W^{[2]T} dz^{[2]} * g^{[1]\prime}(z^{[1]})$$

$$dW^{[1]} = dz^{[1]} x^T$$

$$db^{[1]} = dz^{[1]}$$

In these equations, dz actually means:

$$dz = \frac{\delta L}{\delta z} = \frac{\delta L}{\delta a}\frac{\delta a}{\delta z}$$

Sources:
- Coursera