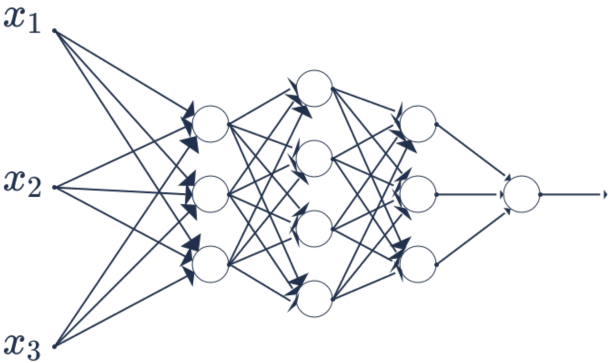
MInDS @ MINES

Deep Learning

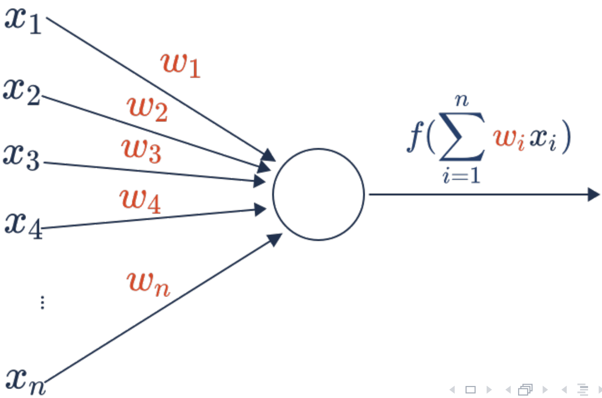
How Data Scientists become magicians

Deep Learning

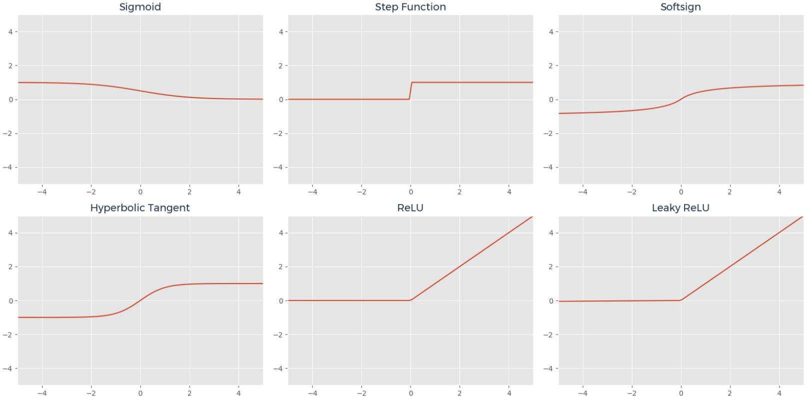


Neuron

$$\mathbf{y} = f(\sum_{i=1}^{n} w_i x_i) = f(\mathbf{w}^T \mathbf{x}),$$

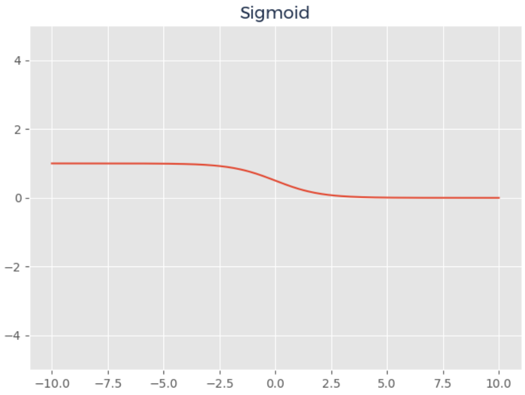


Activation Functions



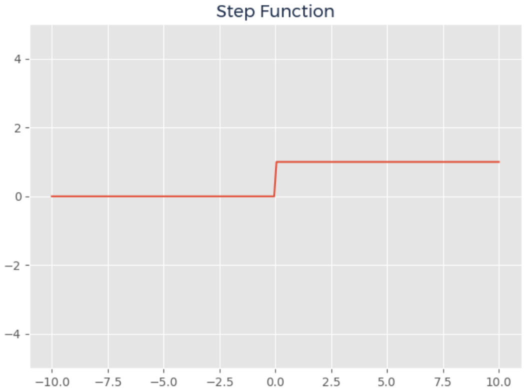
Sigmoid

$$f_{\text{sigmoid}}(x) = \frac{1}{1+e^{-x}}, \quad f'_{\text{sigmoid}}(x) = f_{\text{sigmoid}}(x)(1 - f_{\text{sigmoid}}(x)).$$



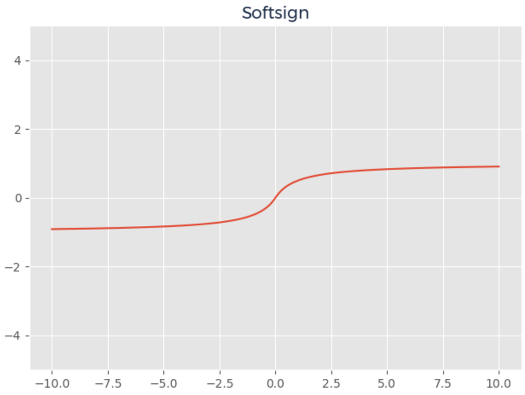
Step

$$f_{\text{step}}(x) = \begin{cases} 0 & x < 0, \\ 1 & x \geq 0, \end{cases} \qquad f'_{\text{step}}(x) = \begin{cases} 0 & x \neq 0, \\ \text{undefined} & x = 0. \end{cases}$$



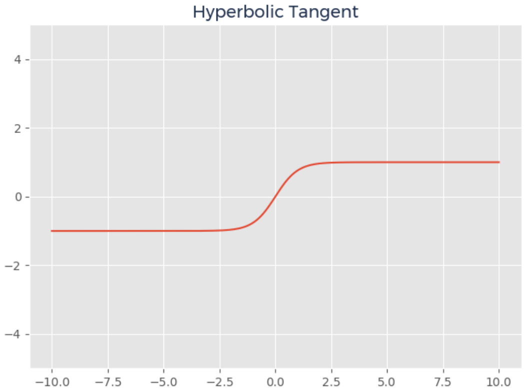
Softsign

$$f_{\text{softsign}}(x) = \frac{x}{1+|x|}, \qquad f'_{\text{softsign}}(x) = \frac{x}{(1+|x|)^2}.$$



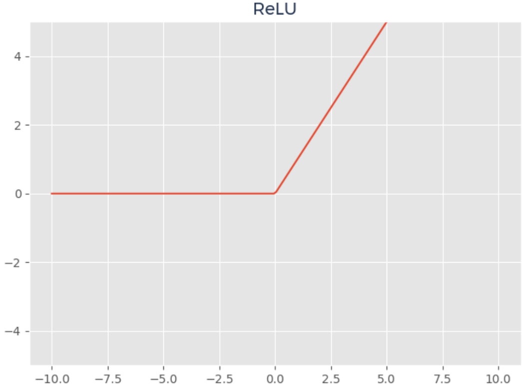
Hyperbolic Tangent

$$f_{\text{hyp. tangent}}(x) = \frac{e^x - e^{-x}}{e^x - e^{-x}}, \quad f'_{\text{hyp. tangent}}(x) = 1 - f_{\text{hyp. tangent}}(x)^2.$$



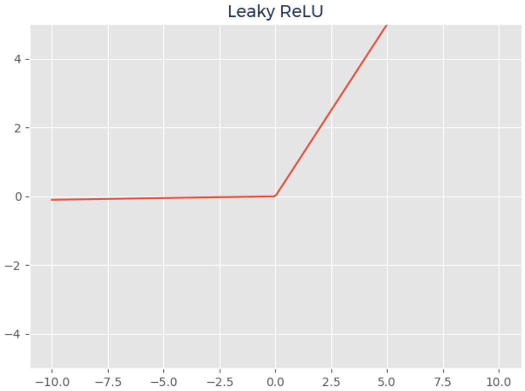
ReLU

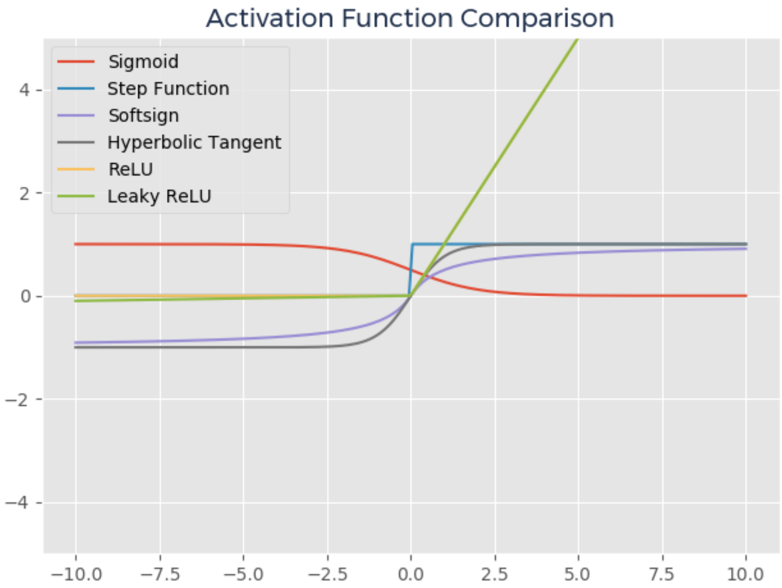
$$f_{\text{ReLU}}(x) = \max(0, x), \qquad f'_{\text{ReLU}}(x) = \begin{cases} 0 & x < 0, \\ 1 & x \geq 0 \end{cases}.$$



Leaky ReLU

$$f_{\text{Leaky ReLU}}(x) = \begin{cases} 0.01x & x < 0, \\ x & x \geq 0, \end{cases} \quad f'_{\text{Leaky ReLU}}(x) = \begin{cases} 0.01 & x \neq 0, \\ 1 & x = 0, \end{cases}.$$




Activation Function Comparison
Sigmoid
Step Function
Softsign
Hyperbolic Tangent
ReLU
Leaky ReLU
4
2
0
−2
−4
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0

Training - Cost / Error Function

$$\epsilon_i = ||\hat{y}_i - y_i||_2.$$

Training - Gradient Descent

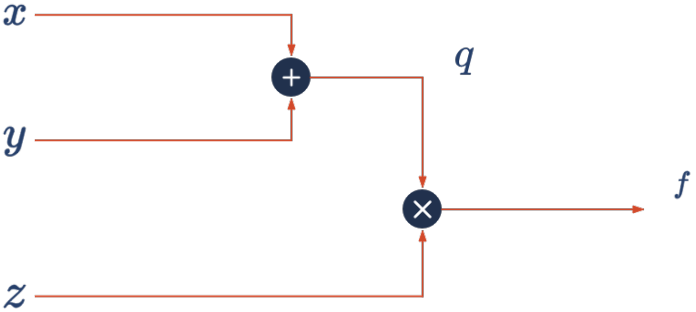
$$\mathbf{w}_{t+1} = \mathbf{w}_t - \alpha \nabla \epsilon$$

Training - Chain Rule

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} \times \frac{\partial y}{\partial x}.$$

Backpropagation Example

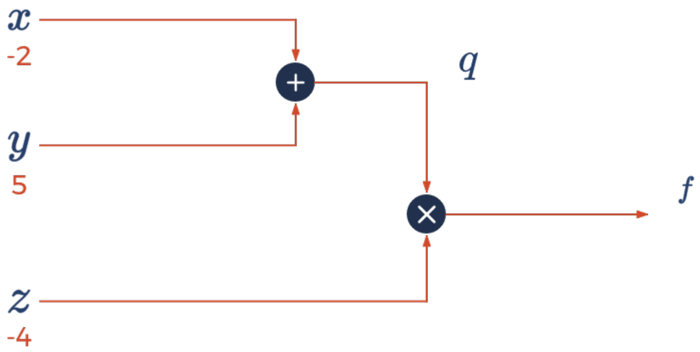
$$f(x, y, z) = (x + y)z$$



Backpropagation Example

$x = -2, y = 5, z = -4$

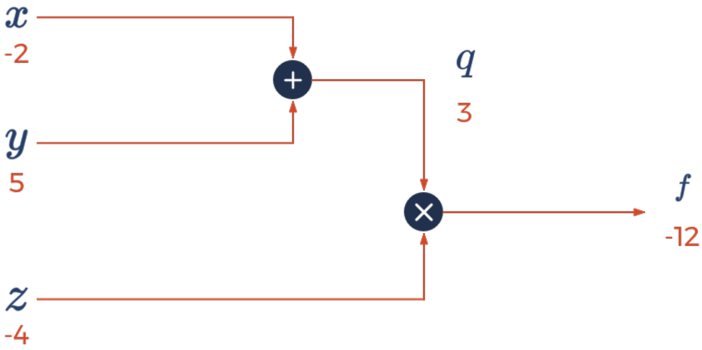
$$f(x, y, z) = (x + y)z$$



Backpropagation Example

$$f(x, y, z) = (x + y)z$$

$x = -2, y = 5, z = -4$



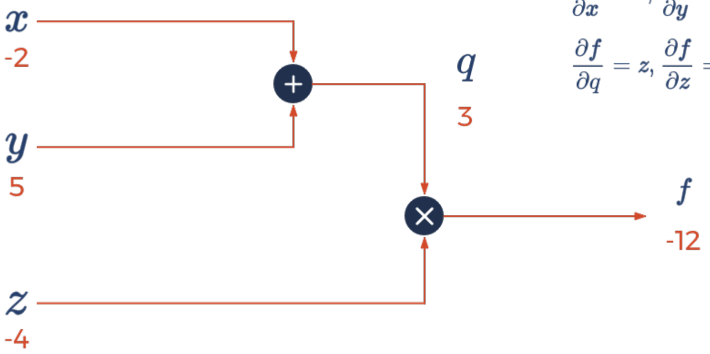
Backpropagation Example

$x = -2, y = 5, z = -4$

$$f(x, y, z) = (x + y)z$$

$$\frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1.$$

$$\frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q,$$



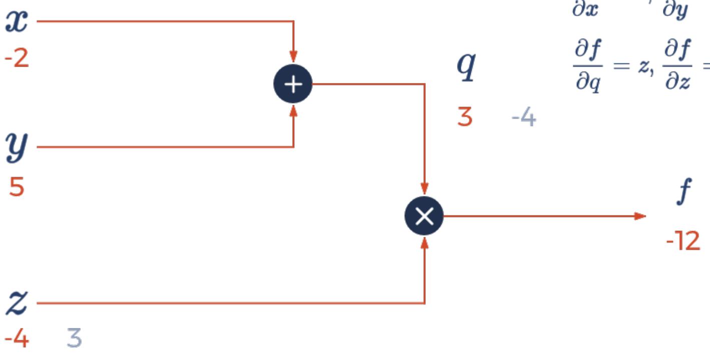
Backpropagation Example

$x = -2, y = 5, z = -4$

$$f(x, y, z) = (x + y)z$$

$$\frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1.$$

$$\frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q,$$



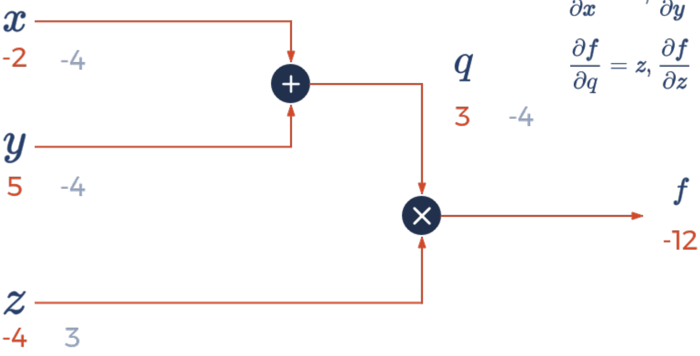
Backpropagation Example

$$f(x, y, z) = (x + y)z$$

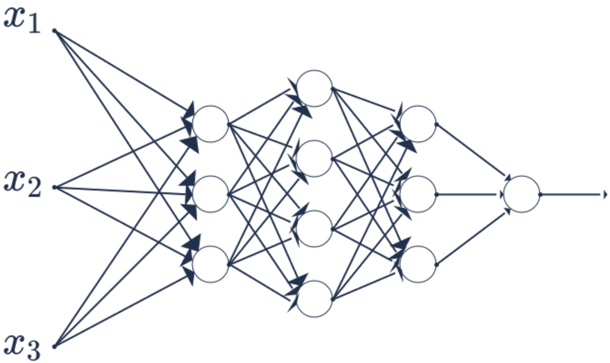
$$x = -2, y = 5, z = -4$$

$$\frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1.$$

$$\frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q,$$



Feed Forward Neural Network

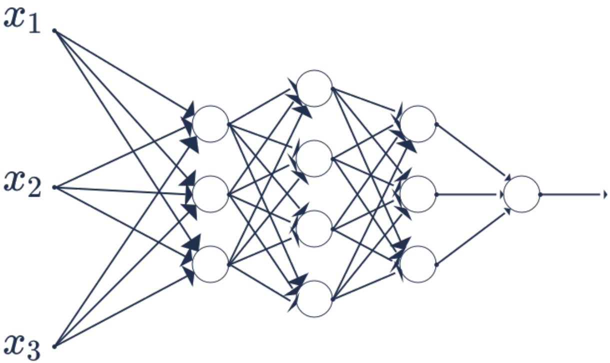


Overfitting - Regularization

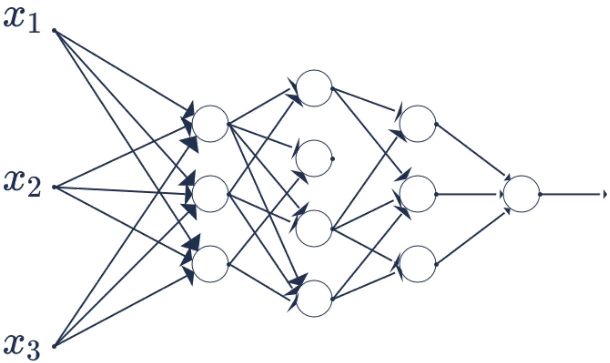
$$\epsilon = ||\hat{\mathbf{y}} - \mathbf{y}||_2.$$

$$\epsilon = ||\hat{\mathbf{y}} - \mathbf{y}||_2 + ||\mathbf{w}||_p^q.$$

Overfitting - Dropout



Overfitting - Dropout



TensorFlow Playground

Demo

Questions