

# Machine Learning

Réseaux de Neurones

---

$$\sigma \left( \begin{bmatrix} x_1 & x_2 & \dots & x_d \end{bmatrix} * \begin{bmatrix} ?_{11} & ?_{12} & \dots & ?_{1n} \\ ?_{21} & ?_{22} & \dots & ?_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ?_{d1} & ?_{d2} & \dots & ?_{dn} \end{bmatrix} \right) = \begin{bmatrix} o_1 & o_2 & \dots & o_n \end{bmatrix}$$

où :

$X$  est une donnée en entrée de dimension  $\mathbf{d}$ ,

$?_{ij}$  sont les paramètres à trouver des  $\mathbf{n}$  neurones de notre modèle.

$\sigma$  la fonction d'activation et

$O$  la sortie du réseau

Calcul de l'erreur

$$\frac{1}{2} \left( \begin{bmatrix} o_1^* & o_2^* & \dots & o_n^* \end{bmatrix} - \begin{bmatrix} o_1 & o_2 & \dots & o_n \end{bmatrix} \right)^2 = \begin{bmatrix} e_1 & e_2 & \dots & e_n \end{bmatrix}$$

où :

$O^*$  représente la sortie attendue du réseau,

$O$  la sortie du réseau et

$E$  l'erreur commise par chaque neurone de sortie.

Mise à jour des poids (**backward**)

$$\Delta w_i = -\gamma * \frac{\partial E}{\partial w_i}$$

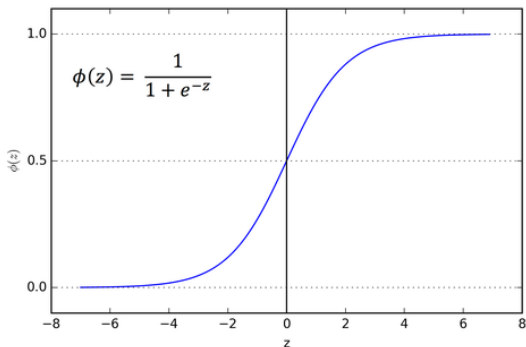
où :

$\Delta w_i$  est l'update du paramètre  $w_i$  et

$\gamma$  est un méta-paramètre du modèle (learning rate)

- Sigmoidé
- Tanh
- Softmax
- ReLU
- ...

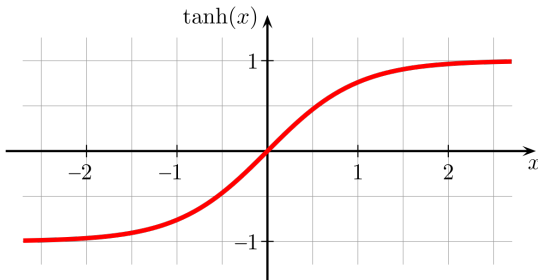
## Sigmoïde



$$\frac{\partial \phi(x)}{\partial x} = \phi(x) * (1 - \phi(x))$$

# Réseaux de neurones : Fonction d'activation $\sigma$

$$\tanh(x) = \frac{1 - \exp -2 * x}{1 + \exp -2 * x}$$



$$\frac{\partial \tanh(x)}{\partial x} = 1 - \tanh^2(x)$$

$$\text{Softmax}(x_j) = \frac{\exp x_j}{\sum_{i=1}^n \exp x_i}$$

donc :

$$\sum_{j=1}^n \text{Softmax}(x_j) = 1$$

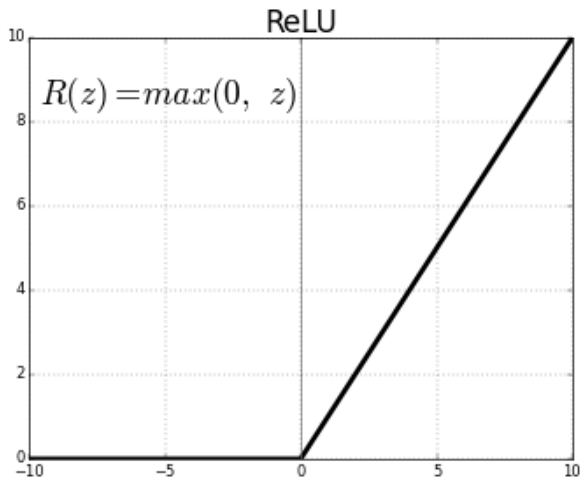
dérivée (ou jacobien car le softmax est une fonction de  $\mathbb{R}^n \rightarrow \mathbb{R}^n$ ) :

$$D_j S_i = S_i(\delta_{ij} - S_j)$$

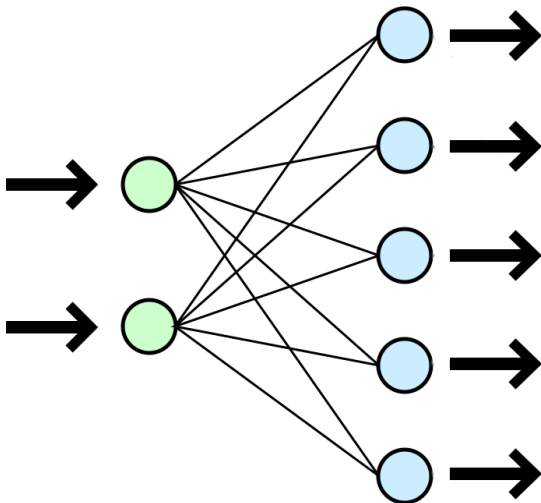
où  $D_j S_i$  est la dérivée partielle de la  $i$ -ième sortie par rapport à la  $j$ -ième entrée

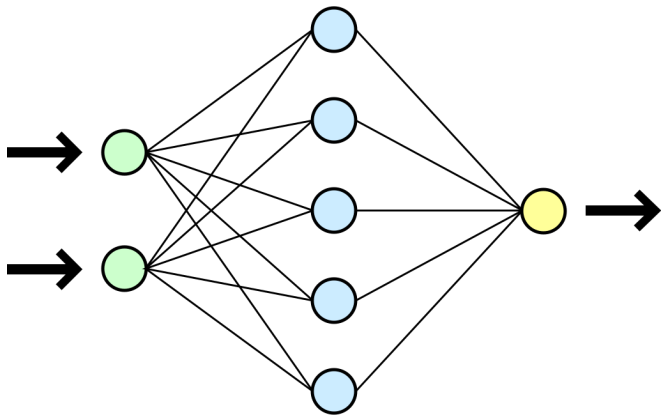
$\delta_{ij}$  est le delta de Kronecker



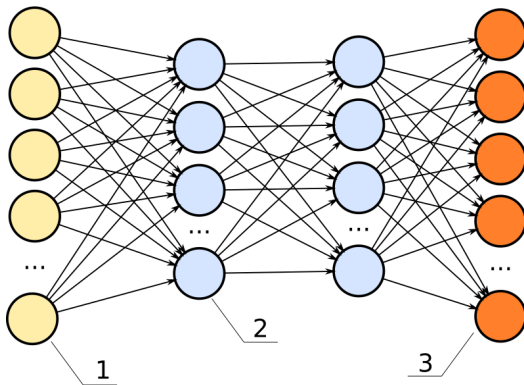


# Réseau de Neurones





# Réseau de Neurones



# Réseau de Neurones

