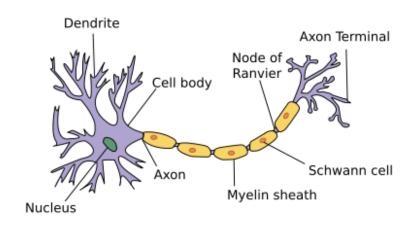
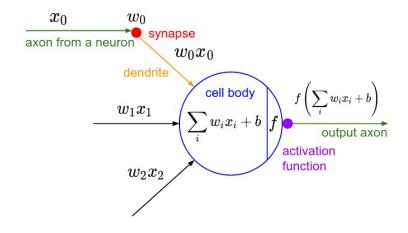
Machine Learning



1



$$\sigma \left(\begin{bmatrix} x_1 & x_2 & \dots & x_d \end{bmatrix} * \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{d1} & w_{d2} & \dots & w_{dn} \end{bmatrix} + \begin{bmatrix} b_1 & b_2 & \dots & b_n \end{bmatrix} \right)$$

$$= \begin{bmatrix} o_1 & o_2 & \dots & o_n \end{bmatrix}$$

où:

X est une donnée en entrée de dimension d,

w et b sont les paramètres à trouver des \mathbf{n} neurones de notre modèle.

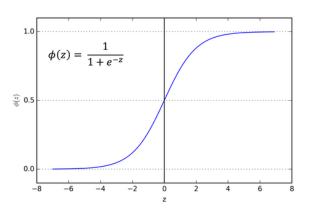
 σ la fonction d'activation et

O la sortie du réseau

3

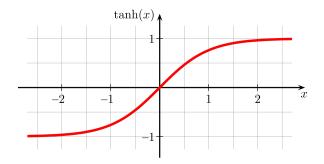
- Sigmoïde
- Tanh
- Softmax
- ReLU
- ...

Sigmoïde



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

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



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

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

donc:

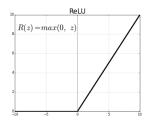
$$\sum_{j=1}^{n} Softmax(x_j) = 1$$

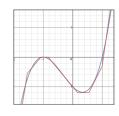
dérivée (ou jacobien car le softmax est une fonction de $\mathbb{R}^n o \mathbb{R}^n$) :

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

où D_jS_i est la dérivée partielle de la i-ième sortie par rapport à la j-ième entrée

 δ_{ii} est le delta de Kronecker





$$\begin{split} &n_1(x) = Retu(-5x-7.7) \\ &n_2(x) = Retu(-1.2x-1.3) \\ &n_3(x) = Retu(1.2x+1) \\ &n_4(x) = Retu(1.2x-2) \\ &n_5(x) = Retu(2x-1.1) \\ &n_6(x) = Retu(5x-5) \end{split}$$

$$Z(x) = -n_1(x) - n_2(x) - n_3(x)$$

 $+ n_4(x) + n_5(x) + n_6(x)$

