

Calcul du gradient des poids d'un neurone

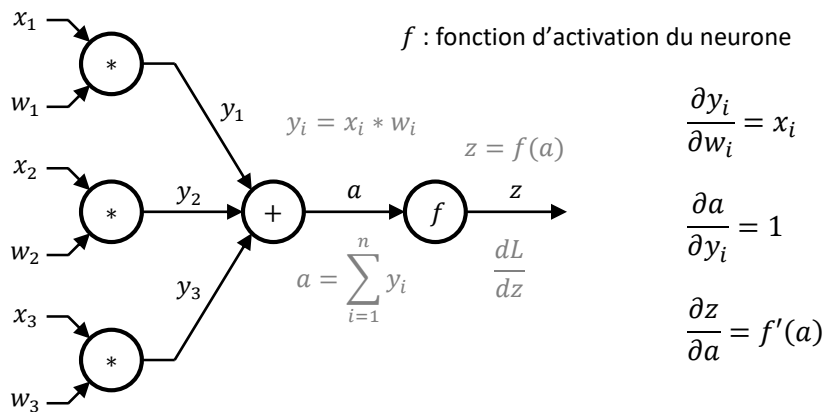
Objectif : Trouver le gradient de chaque poids d'un neurone, c'est-à-dire calculer $\frac{\partial L}{\partial w_i}$

Schéma des 'sous-calculs' d'un neurone :

x_i : valeurs en entrée du neurone

w_i : poids appliqués à chaque entrée correspondante

f : fonction d'activation du neurone



Calculs :

$\frac{\partial L}{\partial z}$ est obtenue grâce au calcul de l'erreur avec la fonction Loss ou en récupérant la valeur $\frac{\partial L}{\partial x_i}$ obtenue depuis la couche précédente (dans l'ordre de la rétropropagation).

Grâce au 'Théorème de dérivation des fonctions composées' :

$$\frac{\partial L}{\partial w_i} = \frac{\partial L}{\partial y_i} \cdot \frac{\partial y_i}{\partial w_i} = \frac{\partial L}{\partial y_i} \cdot x_i$$

Or :

$$\frac{\partial L}{\partial y_i} = \frac{\partial L}{\partial a} \cdot \frac{\partial a}{\partial y_i} = \frac{\partial L}{\partial a}$$

et aussi :

$$\frac{\partial L}{\partial a} = \frac{\partial L}{\partial z} \cdot \frac{\partial z}{\partial a} = \frac{\partial L}{\partial z} \cdot f'(a)$$

Donc :

$$\boxed{\frac{\partial L}{\partial w_i} = \frac{\partial L}{\partial z} \cdot f'(a) \cdot x_i}$$