

Les Réseaux de Neurones

Philibert PAPPENS, Tristan POIDATZ, Mathurin PETIT

April 25, 2022

Contents

1 Définitions	1
1.1 Neurone	1
1.2 Réseau de Neurones	1
1.3 Fonction d'activation	1
2 Notations	2
2.1 Poids	2
2.2 Activation intermédiaire	2
2.3 Biais	2
2.4 Fonction d'activation	2
2.5 Activation	2
2.6 Matrice d'activation	3
2.7 Matrice d'activations intermédiaires	3
2.8 Matrice des poids	3
2.9 Matrice gradient	3
2.10 Matrice δ	3
2.11 Produit de Hadamard	3

1 Définitions

1.1 Neurone

Un Neurone est une représentation mathématique et informatique d'un neurone biologique. Il contient en général plusieurs entrées et une seule sortie. Mathématiquement, un neurone est une fonction à plusieurs variables et à valeurs réelles.

1.2 Réseau de Neurones

Un réseau de Neurones est une association de neurones pour accomplir des tâches arbitrairement complexes.

1.3 Fonction d'activation

Une fonction d'activation (souvent notée σ) est une fonction de $\mathbb{R}^{\mathbb{R}}$ dont son calcul et celui de sa dérivée est peu couteux en temps. Choisir des fonctions d'activations non-linéaires permet au réseau de créer des comportements plus complexes.

2 Notations

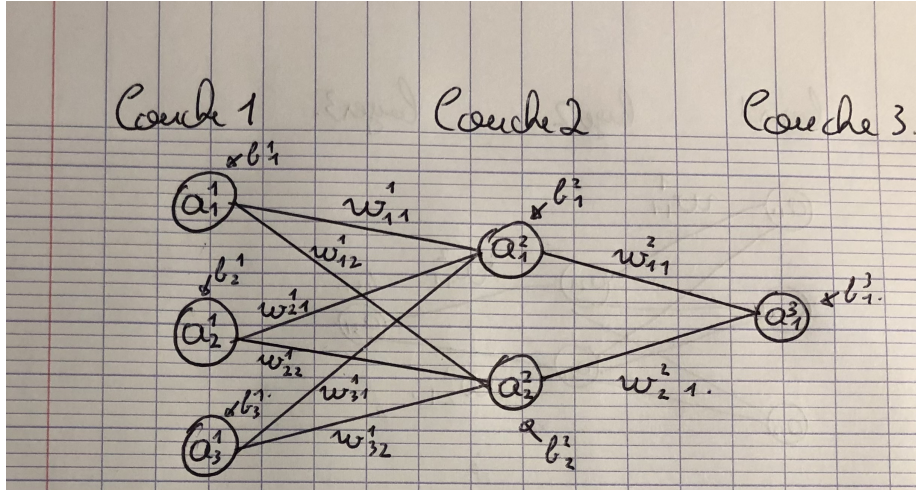


Figure 1: Réseau simple à trois couches

2.1 Poids

$$w_{i,j}^{\ell}$$

est le poids reliant le neurone i de la couche ℓ au neurone j de la couche $\ell + 1$

2.2 Activation intermédiaire

$$z_i^{\ell}$$

est la valeur de la somme pondérée avant passage dans la fonction d'activation.

2.3 Biais

$$b_i^{\ell}$$

est la valeur du biais i de la couche ℓ .

2.4 Fonction d'activation

$$\sigma^{\ell} :$$

est la fonction d'activation permettant de passer de la couche $\ell - 1$ à la couche ℓ .

2.5 Activation

$$a_i^{\ell}$$

est la valeur du neurone i de la couche ℓ . Elle vaut $\sigma(z_i^{\ell})$

2.6 Matrice d'activation

$$A_\ell = (a_1^\ell \quad a_2^\ell \quad \dots \quad a_n^\ell)$$

2.7 Matrice d'activations intermédiaires

$$Z_\ell = (z_1^\ell \quad z_2^\ell \quad \dots \quad z_n^\ell)$$

2.8 Matrice des poids

$$W_\ell = \begin{pmatrix} w_{1,1}^\ell & w_{1,2}^\ell & \dots & w_{1,k}^\ell \\ w_{2,1}^\ell & w_{2,2}^\ell & \dots & w_{2,k}^\ell \\ \dots & \dots & \dots & \dots \\ w_{n,1}^\ell & w_{n,2}^\ell & \dots & w_{n,k}^\ell \end{pmatrix}$$

Où n est le nombre de neurones dans la couche ℓ et k dans la couche $\ell + 1$

2.9 Matrice gradient

Pour une matrice M_ℓ , on note ∇_M^ℓ la matrice des dérivées partielles par rapport à chacun des coefficients. Par exemple:

$$\nabla_A^\ell = \begin{pmatrix} \frac{\partial}{\partial a_1^\ell} & \frac{\partial}{\partial a_2^\ell} & \dots & \frac{\partial}{\partial a_n^\ell} \end{pmatrix}$$

2.10 Matrice δ

On pose $\delta^\ell = \nabla_Z^\ell C$

2.11 Produit de Hadamard

L'opérateur \odot permet de définir le produit de Hadamard: le produit matriciel terme à terme:

$$\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \odot \begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix} = \begin{pmatrix} a_{1,1} \cdot b_{1,1} & a_{1,2} \cdot b_{1,2} \\ a_{2,1} \cdot b_{2,1} & a_{2,2} \cdot b_{2,2} \end{pmatrix}$$