Fonction d’activation

Dans les réseaux de neurones, une fonction d'activation n'est rien d'autre qu'une fonction qui sera dans la suite associée à un neurone.

Cette fonction d'activation va avant tout nous permettre de modifier de manière non linéaire les données. Chaque neurone va appliquer sa fonction d'activation sur les données et les résultats obtenus après de la fonction d'activation vont être différente pour chaque neurone car chaque neurone possède des poids différents.

Nous allons étudier plusieurs exemples de fonctions d’activation. A une entrée , on associe une sortie .

Dans ce documents nous allons voir les fonctions d’activation suivantes :

* Heaviside
* Relu
* Tahn
* Sigmoïde

Mais il en existe bien d’autre comme:

* ELU
* SELU
* Softsign
* Softplus
* Softmax

# Fonctions marche Heaviside

La fonction Heaviside est la fonction marche d’escalier définie par la formule suivante :

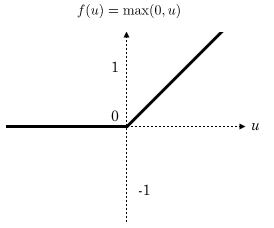
# 

La fonction Heaviside a pour avantage d’être très simple et de séparer clairement le résultat en deux catégories 0 ou 1. Mais ceci pour être gênant car il n’y a pas de nuance possible.

De plus la fonction n’est pas dérivable en 0 ni continue et ceci pourrait être gênant.

# Fonction Relu

La fonction d’activation ReLU et de nos jours la fonction d’activation la plus répandue dans un réseau de neurones.



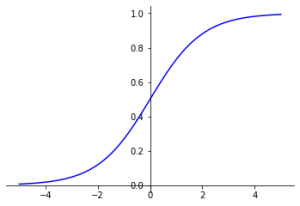
Cette fonction est définie par la formule suivante :

La fonction d’activation ReLU donnera les valeurs positives mais si une valeur est négative elle lui attribuera la valeur .

Attention cette fonction n’est pas dérivable en 0 (sauf dans le cas ou a = b)

# Sigmoïde

La fonction sigmoïde est définie par :



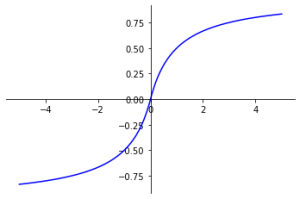
La fonction sigmoïde donne la valeur entre et . Elle est très utilisée pour faire de la classification binaire, quand un modèle doit déterminer seulement deux labels.

De plus :

* La fonction σ est strictement croissante.
* La limite en −∞ est 0, la limite en +∞ en
* La fonction σ est continue et dérivable :
* La tangente en a pour équation :

# Fonction Tanh

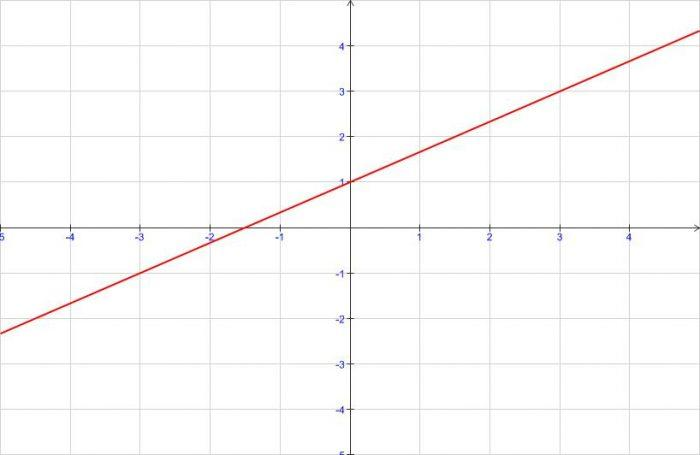
La fonction Tanh est la fonction de la tangente hyperbolique. Elle est définie par :



Elle est similaire à la fonction sigmoïde mais elle ne va pas varier entre et mais plutôt entre et . L’avantage avec cette fonction est que les entrées négatives seront vraiment négatives. Tandis qu’avec la fonction sigmoid, les valeurs négatives peuvent être confondus avec les valeurs proches de 0.

# Fonction affine

la fonction affine est définie par :



Les fonctions affines seront à la base des réseaux de neurones. Elles sont simples.

Si et alors . C’est la fonction identité : la sortie est égale à l’entrée.

Si La sortie est proportionnelle à l’entrée.

Le principal inconvénient est le suivant :

Si et sont deux fonctions affines alors la composition est encore une fonction affine. Ainsi, avec des fonctions d’activation qui sont affines, par composition on n’obtiendra que des fonctions affines, ce qui n’est pas assez riche pour les problèmes de classification.