

Note de synthèse

Dell'Aiera Clément, Prévosteau Clément, David Wahiche

Ce projet a pour objectif de présenter une introduction au réseau de neurones artificiels, et à certaines nouvelles méthodes qui sont récemment appliquées en ce domaine, à savoir les méthodes de la géométrie de l'information, ainsi que l'apprentissage profond, ou deep learning.

Les réseaux de neurones peuvent être rapidement présentés comme des modèles de régression. Ils consistent en plusieurs couches de neurones empilées les une au-dessus des autres, reliées par des connexions dont la réponse dépend d'un poids et d'une fonction d'activation. Ces derniers sont les paramètres du modèle dont l'expérimentateur va faire varier la valeur en entraînant le réseau sur une base de données via un apprentissage supervisé.

Formellement, un réseau de neurones est la donnée d'un entier N , le nombre de couches, de $N-1$ matrices de poids W_j , $j = 1, N-1$ et d'autant de fonctions d'activations ϕ_j . Si $x \in D$. Alors, si l'entrée est une donnée $x \in \mathcal{D}$, le réseau agit récursivement :

$$\begin{cases} x_0 = x \\ x_{j+1} = f_j(W_j x_j) \end{cases}$$

Dans la formule ci-dessus, la fonction d'activation est appliquée à chaque terme du vecteur $W_j x_j$. Les fonctions d'activation sont à choisir parmi celles connues des praticiens : *tanh*, sigmoïde,...

Références

- [1] Shun ichi Amari. Information geometry of the em and em algorithm for neural networks. *Neural Networks*, 8(9) :1379–1408, 1994.
- [2] Shun ichi Amari. Natural gradient works efficiently in learning. *Neural Computation*, 10(2) :251–276, 1998.
- [3] Yann Ollivier. Riemannian metrics for neural networks. *Preprint*, 2013.
- [4] Hiroshi Nagaoka Shun-ichi Amari. *Methods of information geometry*. Oxford University Press, 1993.