



Chapitre 5: Neurone Formel

Ibrahima Sy

Ph.d Cand. in Machine Learning

Université Cheikh Anta Diop de Dakar(UCAD)
Master Modélisation Statistique et Informatique(MSI)
Faculté des Sciences et Techniques (FST)

May 19, 2021

Plan

Présentation

Fonctionnement de L'algorithme

Règle d'apprentissage

References

Présentation

Un neurone formel est une représentation mathématique et informatique d'un neurone biologique. Le neurone formel possède généralement plusieurs entrées et une sortie qui correspondent respectivement aux dendrites et au cône d'émergence du neurone biologique (point de départ de l'axone). Les actions excitatrices et inhibitrices des synapses sont représentées, la plupart du temps, par des coefficients numériques (les poids synaptiques) associés aux entrées. Les valeurs numériques de ces coefficients sont ajustées dans une phase d'apprentissage. Dans sa version la plus simple, un neurone formel calcule la somme pondérée des entrées reçues, puis applique à cette valeur une fonction d'activation, généralement non linéaire. La valeur finale obtenue est la sortie du neurone.

Présentation

Le neurone formel est l'unité élémentaire des réseaux de neurones artificiels dans lesquels il est associé à ses semblables pour calculer des fonctions arbitrairement complexes, utilisées pour diverses applications en intelligence artificielle.

Régression Logistique: Cas deux classes

- ▶ Plutôt que de prédire une classe, prédire une probabilité d'appartenir à la classe 1

$$\mathbb{P}(y_t = 1|x) = h_w(x_t) = \text{logistic}(w.x) = \frac{1}{1+\exp(-w.x)}$$

- ▶ Choisir la classe la plus probable selon le modèle
 - ▶ $h_w(x_t) \geq 0.5$ choisir la classe 1
 - ▶ sinon, choisir la classe 0

Régression Logistique: Cas deux classes

1. base des réseaux de neurones
2. Entré : Vecteurs $x \in \mathbb{R}^m$
3. Neurone de sortie : $\hat{y} \in \mathbb{R}$
4. Transformation affine : $s = \omega^\top x + b$
5. Transformation non linéaire (fonction activation) : $f_w(s) = \hat{y}$

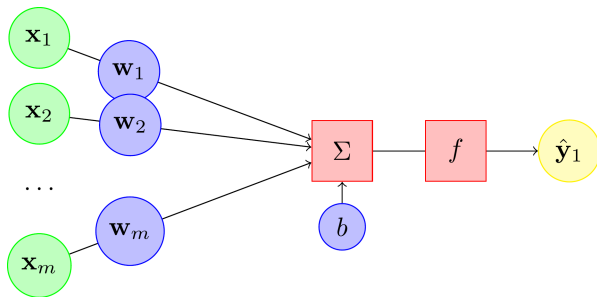
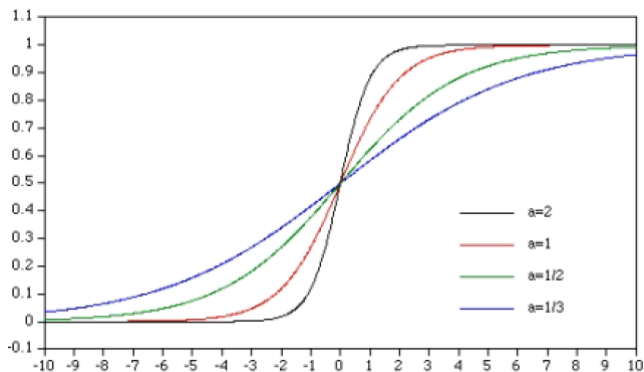


Figure 1: Neurone formel

Fonction Activation

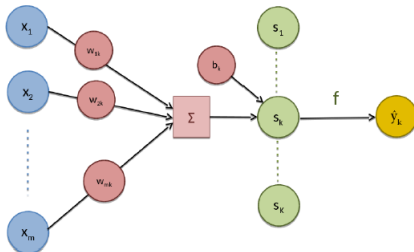
► Fonction Sigmoidé

$$\sigma(z) = \frac{1}{1 + \exp(-az)}$$



Régression Logistique: Cas multiclass

1. Pour des une classification en K classes
2. Entré : Vecteurs $x \in \mathbb{R}^m$
3. Neurone de sortie : $\hat{y} \in \mathbb{R}^k$
4. Transformation affine : $s_k = \omega_k^\top x + b_k$ avec $\omega_k = (\omega_{1k}, \omega_{2k}, \dots, \omega_{mk})$
5. Transformation non linéaire (fonction activation) : $f(s_k) = \hat{y}_k$
6. fonction soft-max: $\hat{y}_k = f(s_k) = \frac{\exp s_k}{\sum_{k'=1}^K \exp s'_k}$



Règle d'apprentissage

- Pour obtenir une règle d'apprentissage, on définit d'abord une perte

$$Loss(y_t, h_w(x_t)) = -y_t \log(h_w(x_t)) - (1 - y_t) \log(1 - h_w(x_t))$$

- Si $y_t = 1$, on souhaite maximiser la probabilité $\mathbb{P}(y_t = 1|x) = h_w(x_t)$
- Si $y_t = 0$, on souhaite maximiser la probabilité $\mathbb{P}(y_t = 0|x) = 1 - h_w(x_t)$

- On dérive la règle d'apprentissage comme une descente de gradient

$$w_i \leftarrow w_i - \alpha \frac{\partial}{\partial w_i} \text{Loss}(y_t, h_{\mathbf{w}}(\mathbf{x}_t)) \quad \forall i$$

ce qui donne

$$w_i \leftarrow w_i + \alpha (y_t - h_{\mathbf{w}}(\mathbf{x}_t)) x_{t,i} \quad \forall i$$

- La règle est donc la même que pour le Perceptron, mais la définition de

References I

- ▶ **Hugo Larochelle**, Professeur associé, Université de Montréal, Google
- ▶ **Pierre-Marc Jodoin**, Professeur titulaire Université Sherbrooke
- ▶ Bayesian Reasoning and Machine Learning de David Barber
- ▶ The Elements of Statistical Learning de Trevor Hastie,
- ▶ Robert Tibshirani et Jerome Friedman
- ▶ Information Theory, Inference, and Learning Algorithms de David J.C. MacKay
- ▶ Convex Optimization de Stephen Boyd et Lieven Vandenberghe
- ▶ Natural Image Statistics de Aapo Hyvärinen, Jarmo Hurri et Patrik O. Hoyer
- ▶ The Quest for Artificial Intelligence - A History of Ideas and Achievements de Nils J. Nilsson
- ▶ Gaussian Processes for Machine Learning de Carl Edward Rasmussen et Christopher K. I. Williams
- ▶ Introduction to Information Retrieval de Christopher D. Manning, Prabhakar Raghavan et Hinrich Schütze