

Atelier 3

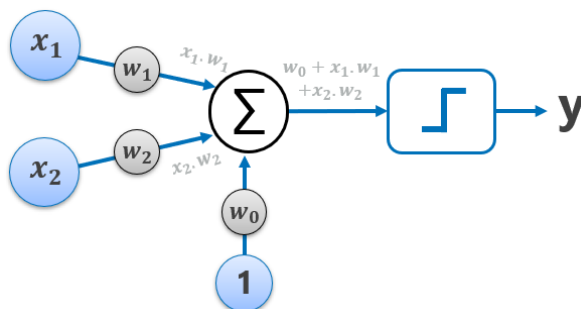
Introduction au deep learning

Objectifs

- Appliquer un algorithme de deep learning à une problématique simple
- Créer un modèle d'intelligence artificielle basé sur le concept du perceptron

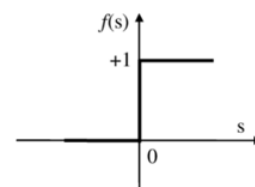
Rappels théoriques

Représentation du perceptron avec 2 entrées



Le y_k estimé est donnée par la formule suivante :

$$y = f(w_0 + x_1 \cdot w_1 + x_2 \cdot w_2) = \begin{cases} 1 & \text{si } \sum_{i=1}^n \omega_i x_i \geq \theta \\ 0 & \text{sinon} \end{cases}$$



Fonction de seuil

L'erreur est déterminée par :

$$(y_{1 \text{ estimé}} - y_{1 \text{ réel}})(w_0 + x_{1,1} \cdot w_1 + x_{1,2} \cdot w_2)$$

Les poids sont mis à jour à partir de la formule suivante :

$$w_i^{k+1} = w_i^k - \alpha(y_{i \text{ estimé}} - y_{i \text{ réel}})x_i$$

Problématique

Suite à la découverte de PFAS (polluants) dans l'eau de distribution à Chièvres, la Société Wallonne Des Eaux (SWDE) a décidé d'analyser l'eau circulant dans le réseau à différents endroits de la région en Wallonie.

20 relevés ont été effectués et les résultats qui en découlent sont répertoriés dans le tableau suivant :

Numéro de Relevés	Localisation		Pollution
	Latitude	Longitude	
1	2	6	OUI
2	7	11	OUI
3	7	4	NON
4	10	8	NON
5	1	0	OUI
6	7	5	NON
7	6	12	OUI
8	10	11	NON
9	11	6	NON
10	4	13	OUI
11	7	4	NON
12	3	3	OUI
13	9	2	NON
14	5	8	OUI
15	7	14	OUI
16	5	12	OUI
17	7	4	NON
18	11	13	NON
19	8	4	NON
20	2	2	OUI

La SWDE nous demande de leur fournir un modèle qui va estimer si l'eau qui circule à un autre endroit donné du réseau est polluée ou non.

Le modèle doit être réalisé à partir d'un perceptron simple.

Etape 0 : Phase de préparation des données

Dans cette étape, vous allez préparer les données fournies afin qu'elles puissent être exploitables pour notre modèle.

- Dans la feuille « **DATA** » du classeur « LABO_DL.xlsx », ajoute une colonne « Y estimé » au tableau de données dans laquelle tu indiques 1 si le site est pollué et 0 sinon. Je te conseille d'utiliser la fonction SI() de Excel.
- Visualise ces données dans un graphique (Série 1 uniquement avec les sites pollués et Série 2 uniquement avec les sites non pollués)

Etape 1 : Phase d'initialisation

Tu dois maintenant fixer des valeurs initiales pour les paramètres de notre perceptron.

Utilise les valeurs suivantes des paramètres :

$\alpha =$	0,5
$\omega_0 =$	0
$\omega_1 =$	1
$\omega_2 =$	1

- Ecris l'équation caractéristique du perceptron à ce stade.
- Trace cette équation graphiquement. Idéalement, utilise le même graphique sur lequel se trouve tes données.
- Quelle interprétation graphique peux-tu en déduire ?

Etape 2 : Calcul du $y_{k \text{ estimé}}$

- A partir des valeurs initiales des paramètres, calcule le $y_{k \text{ estimé}}$.
- Calcule la valeur de l'erreur ?

Etape 3 : Mise à jour des poids

- A partir de la première donnée du tableau, calcule la nouvelle valeur des paramètres.
- Ecris l'équation caractéristique du perceptron à ce stade.
- Trace cette équation graphiquement. Idéalement, utilise le même graphique sur lequel se trouve tes données.
- Comment a évolué ton graphique ?

Etape 4 : Itération et calcul des valeurs des paramètres

- Répète l'étape 3 jusqu'à obtenir un résultat satisfaisant.
Arrivée à la dernière donnée disponible du tableau, continue tes itérations en reprenant la première donnée du tableau et ainsi de suite.
- Comment se traduit graphiquement l'évolution des valeurs de paramètres ?

Etape 5 : Test avec des autres valeurs initiales de paramètres (idem pour α)

- Quelle est l'influence des valeurs initiales sur les calculs ?
- Y a-t-il d'autres caractéristiques du perceptron que nous pourrions changer ? Que proposez-vous ?