

TD N°2

Exercice 1 (perceptron simple)

Soit un perceptron dont le vecteur de pondération $(w_0, w_1, w_2)^T = (2, 1, 1)^T$.

1. Tracer sur un diagramme le séparateur linéaire obtenu par ce perceptron et hachurer la surface correspondante à la partie du plan où le perceptron retourne la valeur 1.
2. Lesquels parmi les perceptrons caractérisés par les vecteurs de pondération suivants ont le même hyperplan et qui retourne exactement le même résultat de classification que le perceptron donnée dans 1. ?

(a) $(w_0, w_1, w_2)^T = (1, 0.5, 0.5)^T$

(b) $(w_0, w_1, w_2)^T = (200, 100, 100)^T$

(c) $(w_0, w_1, w_2)^T = (\sqrt{2}, \sqrt{1}, \sqrt{1})^T$

(d) $(w_0, w_1, w_2)^T = (-2, -1, -1)^T$

3. Quelles sont les valeurs des poids w_0, w_1 et w_2 du perceptron dont la frontière de décision est illustrée ci-dessous ? Y a-t-il plusieurs choix possibles pour ces valeurs de poids ? Si les étiquettes de classification (+ ou -) sont inversées, les poids resteront-ils les mêmes ?

Exercice 2 :

1. Appliquer l'algorithme du perceptron jusqu'à convergence sur les données d'apprentissage suivantes :

$(4, 3, 6)^T \in N$

$(2, -2, 3)^T \in P$

$(1, 0, -3)^T \in P$

$(4, 2, 3)^T \in N$

Appliquer l'algorithme sur les données dans l'ordre et de manière cyclique en initialisant le vecteur de pondération $(w_0, w_1, w_2, w_3)^T$ à $(1, 0, 0, 0)^T$. Pour chaque étape de l'apprentissage du perceptron, donner le résultat de la classification et de la mise à jour du vecteur de pondération.

2. Montrer que le problème posé par les données d'apprentissage ci-dessous ne peut pas être résolu avec un seul perceptron. Pour cela, appliquer l'algorithme d'apprentissage du perceptron pour ces données d'apprentissage et commencer par le vecteur de pondération $(1, 0, 0)^T$.

$(1, 1)^T \in P$

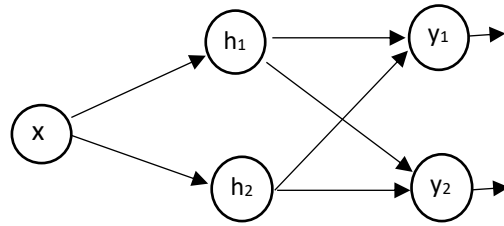
$(1, 0)^T \in N$

$(0, 0)^T \in P$

$(0, 1)^T \in N$

Exercice 3 : (perceptron multicouche-Rétropropagation du gradient)

On considère un réseau à 2 couches de neurones, avec 2 neurones h_1, h_2 dans la couche cachée et 2 neurones y_1, y_2 dans la couche de sortie. L'objectif du réseau est la classification des entrées x en deux classes y_1 ou y_2 .



Tous les neurones possèdent la même fonction d'activation non-linéaire : $f(x) = \frac{1}{1+e^{-x}}$

On suppose les poids tous initialisés à 0. Le taux d'apprentissage $\eta=0.1$.

Question : Présentez au réseau l'élément de la base d'apprentissage $(x, (y_1, y_2)) = (1, (1, 0))$ et le lui faire apprendre par l'algorithme de rétropropagation du gradient, par une itération d'apprentissage. **Donnez les poids W_{ijk} du réseau obtenus.**