

# Intelligence Artificielle - TD<sup>o</sup>4

## Réseaux de neurones artificiels

### Fonctionnement du Perceptron

#### Exercice 1

On considère la base d'apprentissage donnée dans le tableau suivant :

Entrée			Sortie
3	2	1	0
1	1	1	1
1	2	3	1

En utilisant le pas d'apprentissage  $\delta = 0.1$  et l'architecture du perceptron suivante :

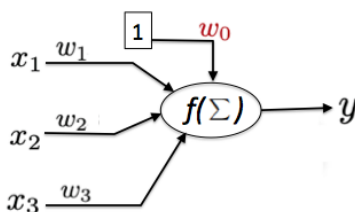


Figure 1: Schéma du perceptron (1)

1. Dérouler manuellement l'algorithme du perceptron en utilisant cette base d'apprentissage
2. Coder l'algorithme d'apprentissage du perceptron en utilisant cette base d'apprentissage

#### Exercice 2

On souhaite construire un perceptron dont la sortie est 1 ou 0. Ce perceptron, associé à un pont, est sensé informer un véhicule donné s'il peut ou pas traverser le pont. Chaque véhicule est caractérisé par deux paramètres : sa longueur en mètre et le bruit en décibel (*DB*) qu'il génère. La fonction d'activation du perceptron est la fonction seuil ( $f(x) = 1$  si  $x \leq 0$  et  $f(x) = 0$  si  $x > 0$ ). L'architecture du perceptron utilisé est la suivante :

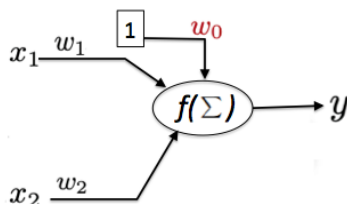


Figure 2: Schéma du perceptron (2)

On dispose du tableau suivant qui donne des informations sur 8 véhicules différents avec l'information : peuvent-ils traverser le pont ou pas.

Véhicule	Caractéristiques		Classe
$V_i$	$x_1$	$x_2$	$C_i$
$V_1$	19	11	1
$V_2$	16	19	1
$V_3$	15	11	0
$V_4$	14	6	0
$V_5$	7	14	0
$V_6$	16	15	1
$V_7$	3	19	0
$V_8$	7	15	0

Si une classe  $C_i = 0$ , le véhicule  $V_i$  correspondant peut traverser le pont et inversement. On considère aussi les valeurs des coefficients synaptiques (poids) :

$$w_0 = -46, w_1 = 2, w_2 = 1$$

1. Avec les poids donnés, le perceptron peut-il remplacer le tableau ? Justifier.
2. Dans le cas contraire, proposer de nouvelles valeurs aux poids en codant l'algorithme d'apprentissage du perceptron. Préciser l'architecture de ce dernier.

### Apprentissage du perceptron

**$S$  : base d'apprentissage  $(x_i, d_i : i = 1, \dots, m)$**   
 $x_i = (x_{i0}, x_{i1}, \dots, x_{in})$   
**On veut déterminer les poids  $(w_0, w_1, \dots, w_n)$**   
**Initialiser aléatoirement le vecteur  $w = (w_0, w_1, \dots, w_n)$**   
**repeat**  
  Prendre un exemple  $(x_j, d_j)$  dans  $S$   
  Calculer la sortie  $o$  du réseau pour l'entrée  $x_j$   
  **Mise à jour des poids**  
  **for**  $i$  allant de 0 à  $n$  **do**  
     $w_{i+1} = w_i + \Delta \times (d_j - o) \times x_j$   
  **end for**  
**until** Convergence

## Fonctionnement du Perceptron Multi-Couches

### Exercice 1

On souhaite modéliser autrement la fonction trigonométrique suivante via les PMC :

$$f(x, y) = (\cos(x) + \sin(x)) \times \cos(y) + (\cos(x) - \sin(x)) \times \sin(y)$$

1. Sans passer par l'étape d'apprentissage, construire un PMC avec une seule couche cachée, des entrées linéaires et une sortie linéaire. Le nombre de cellules cachées, le nombre d'entrées et les fonctions d'activation restent à déterminer.
2. En passant par l'étape d'apprentissage, construire un PMC avec une seule couche cachée, des entrées linéaires et une sortie linéaire. Le nombre de cellules cachées, le nombre d'entrées et les fonctions d'activation restent à déterminer.

## Exercice 2

On considère un système non linéaire ayant deux entrées  $x_1$  et  $x_2$  et une sortie  $y$ . Nous avons construit un *RNA* pour modéliser ce système. Le schéma suivant donne une idée sur le système et le PMC associé :

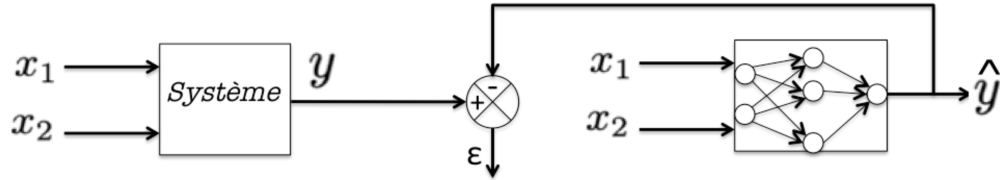


Figure 3: Système MISO identifié par un RNA. )

1. Est-il nécessaire d'avoir une couche cachée dans ce *RNA* ? Justifier.
2. La base d'apprentissage ainsi que les sorties du réseau de neurones après un premier apprentissage sont donnés dans le tableau suivant :

Entrées		Sortie	Sortie RNA
$x_1$	$x_2$	$y$	$\hat{y}$
$v_{x_1}^1$	$v_{x_2}^1$	0.25	0.25
$v_{x_1}^2$	$v_{x_2}^2$	1	1.01
$v_{x_1}^3$	$v_{x_2}^3$	-1	-1
$v_{x_1}^4$	$v_{x_2}^4$	-0.5	-0.5
$v_{x_1}^5$	$v_{x_2}^5$	0.55	0.55
$v_{x_1}^6$	$v_{x_2}^6$	0.25	0.245
$v_{x_1}^7$	$v_{x_2}^7$	-0.75	-0.745
$v_{x_1}^8$	$v_{x_2}^8$	0.55	0.549

Table 1: Base d'apprentissage + Sorties du RNA

Calculer le taux d'apprentissage, qu'en pensez-vous ?

3. La base d'apprentissage ainsi que les sorties du réseau de neurones après un deuxième apprentissage sont donnés dans le tableau suivant :

Entrées		Sortie	Sortie RNA
$x_1$	$x_2$	$y$	$\hat{y}$
$v_{x_1}^1$	$v_{x_2}^1$	0.25	0.25
$v_{x_1}^2$	$v_{x_2}^2$	1	1
$v_{x_1}^3$	$v_{x_2}^3$	-1	-1
$v_{x_1}^4$	$v_{x_2}^4$	-0.5	-0.5
$v_{x_1}^5$	$v_{x_2}^5$	0.55	0.55
$v_{x_1}^6$	$v_{x_2}^6$	0.25	0.25
$v_{x_1}^7$	$v_{x_2}^7$	-0.75	-0.75
$v_{x_1}^8$	$v_{x_2}^8$	0.55	0.55

Table 2: Base d'apprentissage + Sorties du RNA

Quelle est la différence par rapport au premier apprentissage ? Calculer le taux d'apprentissage.

4. Une fois que les paramètres du *RNA* sont figés selon le premier apprentissage, d'autres exemples lui ont été présentés et nous avons obtenu les résultats suivants : Calculer le

Entrées		Sortie	<i>Sortie RNA</i>
$x_1$	$x_2$	$y$	$\hat{y}$
$v_{x_1}^9$	$v_{x_2}^9$	0.4	0.39
$v_{x_1}^{10}$	$v_{x_2}^{10}$	0.3	0.3
$v_{x_1}^{11}$	$v_{x_2}^{11}$	-0.4	-0.385
$v_{x_1}^{12}$	$v_{x_2}^{12}$	0.2	0.189
$v_{x_1}^{13}$	$v_{x_2}^{13}$	-0.1	-0.11
$v_{x_1}^{14}$	$v_{x_2}^{14}$	-0.6	-0.615

taux de généralisation (validation) et le taux de bonnes sorties ? Quel est le lien entre les deux ?

### Exercice 3

Considérons la base de données dont le contenu est composé de descripteurs de tumeurs (entrées) avec les étiquettes relatives (sorties). Ces étiquettes indiquent si la tumeur est maligne (1) ou bénigne (0). On souhaite concevoir un réseau de neurones du type MLP qui puisse donner une indication sur le type de la tumeur en fonction des caractéristiques relatives.

1. Concevoir un *PMC* possédant 9 entrées principales, une couche cachée à 5 neurones cachés et une sortie ( $T$ ). Pour la fonction d'activation au niveau des cellules cachées, on testera les deux fonctions :

(a) La sigmoïde :

$$f(x) = \frac{1}{1 + e^{-x}}$$

(b) La tangente hyperbolique :

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

2. Dans le cas où les résultats précédents ne sont pas satisfaisants avec ce premier réseau à deux cellules cachées, on augmentera progressivement le nombre de cellules cachées.

Pour l'apprentissage, on utilisera 70% d'exemples pour l'apprentissage et 30% d'exemples pour le test de validation.