RESEAUX DE NEURONES avec solution

Exercice 1

L'ensemble d'apprentissage suivant est linéairement séparable :

							classe
$\vec{x_1}$	:	0	0	0	1		0
$\vec{x_2}$	:	0	1	1	1		1
$\vec{x_3}$	:	1	1	0	1		1
$\vec{x_4}$	:	0	0	1	0		0
$\vec{x_5}$	:	0	0	1	1		0
$\vec{x_6}$	:	1	0	0	1		1

Entraînez un perceptron avec cet ensemble et la procédure de correction d'erreur. Le vecteur de poids du perceptron est un vecteur à cinq dimensions (la première composante pour le seuil).

Commencez avec w = (0; 0; 0; 0; 0). Ne dépassez pas 6 itérations.
Il est conseillé de présenter les vecteurs un par un dans l'ordre. (η=1)

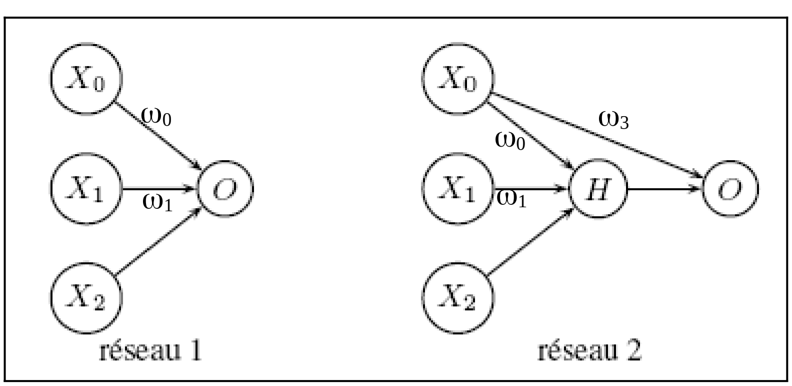
SOLUTION

	W0	**W1**	**W2**	**W3**	**W4**	**Entrée**	**W.X**	**o**	**c**	**c-o**
X1	0	0	0	0	0	10001	0	0	0	0
X2	0	0	0	0	0	10111	0	0	1	1
X3	1	0	1	1	1	11101	3	1	1	0
X4	1	0	1	1	1	10010	2	1	0	-1
X5	0	0	1	0	1	10011	1	1	0	-1
X6	-1	0	1	-1	0	11001	-1	0	1	1
	0	**1**	**1**	**-1**	**1**					

W=(0,1,1,-1,1)

Exercice 2

Considérons les deux réseaux :



ω_4

ω_2

$X_0 = 1$, la fonction Erreur est **J= ½(D-O)²**, D : sortie désirée, O : sortie calculée

On veut étudier le problème XOR (apprendre la fonction XOR)

X_1	X_2	D
-1	-1	-1
-1	1	1
1	-1	1
1	1	-1

Le problème XOR (-1 signifie faux et 1 signifie vrai)

0. Représenter XOR dans le plan (X_1, X_2) : rond pour la classe -1 et croix pour la classe 1.

Posons A est la somme pondérée :

$$A = \sum_{i=0}^{2} w_i X_i$$

1. Considérons le réseau 1, avec une sortie linéaire (fonction d'activation est l'identité):

$$O = A = \sum_{i=0}^{2} w_i X_i$$

calculer $\partial J / \partial w_i$

2. Même question en changeant la fonction d'activation de ce premier réseau :

$$O = \frac{1 - e^{-A}}{1 + e^{-A}}$$

calculer $\partial J / \partial w_i$

3. Considérons maintenant le réseau 2, avec les fonctions d'activation suivantes :
H : sortie de la cellule cachée.

$$
\begin{aligned}
H &= \frac{1-e^{-A}}{1+e^{-A}} \\
O &= w_3 X_0 + w_4 H
\end{aligned}
$$

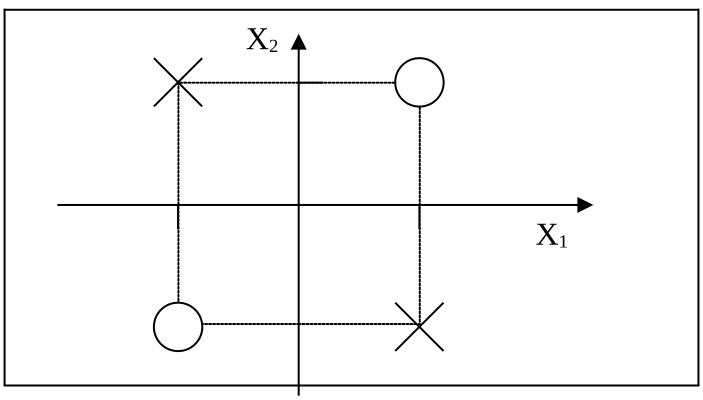
calculer $\partial J/\partial w_i$
pour la couche de sortie et la couche cachée.

4. On considère maintenant les réseaux 1 et 2 avec une fonction d'activation **signe (O = -1 si A < 0 et O=1 si A > 0)** et le problème XOR.

Choix de la structure :
Les réseaux 1 ou 2 sont-ils capables de résoudre ce problème. Si c'est le cas, expliquer pourquoi, sinon un autre réseau du même genre avec 2 cellules dans la couche cachée peut-il résoudre le problème?. Donner les poids dans ce cas.

SOLUTION

0. XOR :



1.

$$\frac{\partial J}{\partial w_i} = \frac{\partial J}{\partial O}\frac{\partial O}{\partial w_i} = (O-D)X_i$$

2.

$$\frac{\partial J}{\partial w_i} = \frac{\partial J}{\partial O}\frac{\partial O}{\partial A}\frac{\partial A}{\partial w_i} = (O-D)\frac{2e^{-A}}{(1+e^{-A})^2}X_i = \frac{1}{2}(O-D)(1-O^2)X_i$$

3.
a) couche de sortie O
les poids : $u_i \in \{w_3, w_4\}$

$lesentr$ées.$Z_i \in$ ¶X_0 , H◊⇔$\mathcal{H}$∫∇⊔⟩⌉∏↕⊣⌋≀⊓⌋⟨⌈⌋⊣⌋⟨é⌉$\mathcal{H}$

$$\frac{\partial J}{\partial u_i} \quad = \quad \frac{\partial J}{\partial O}\frac{\partial O}{\partial u_i} = (O-D)Z_i$$

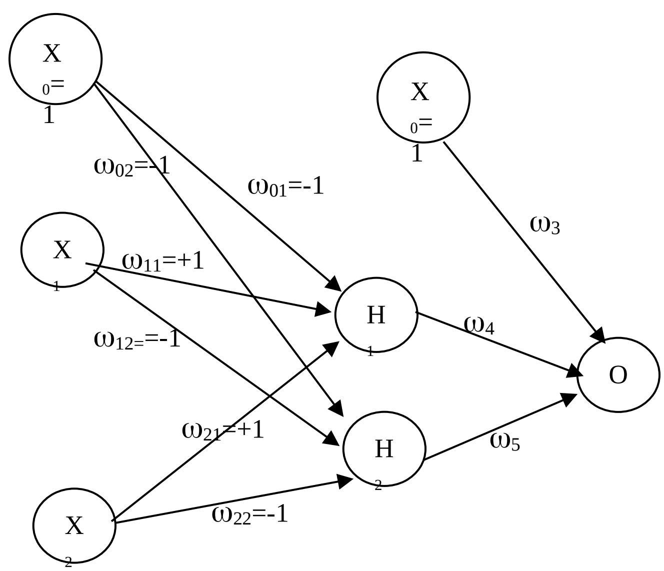
b) pour la couche cachée

$$\frac{\partial J}{\partial w_i} \quad = \quad \frac{\partial J}{\partial O}\frac{\partial O}{\partial H}\frac{\partial H}{\partial A}\frac{\partial A}{\partial w_i} = (O-D)w_4\frac{1}{2}(1-H^2)X_i$$

4.
a)
L'équation du réseau 1 correspond à celle d'un plan dans l'espace ¶X_1‹X_2◊. Si la

fonction de transfert est la fonction signe, le réseau 1 ne permet de résoudre que des problèmes linéairement séparables, ce qui n'est pas le cas du problème XOR.
L'équation du réseau 2 permet seulement de déplacer le plan déterminé par la couche cachée. Ce n'est pas suffisant pour résoudre XOR (on peut même noter que tout problème linéairement séparable résoluble par le réseau 2 peut l'être par le réseau 1 tout seul !)
Pour résoudre XOR, il faut donc combiner 2 frontières de décision, ce qui ne peut se faire qu'en
utilisant 2 neurones dans la couche cachée.

b) Supposons le réseau avec 2 neurones cachés avec les poids affichés dans la figure suivante.



in_H_1= 1*ω_{01} + X_1*ω_{11} + X_2*ω_{21}

$in\_H_2 = 1*\omega_{02} + X_1*\omega_{12} + X_2*\omega_{22}$

Prenons tout simplement comme fonction d'activation la fonction signe la couche cachée et la sortie :

$H_i = f(in\_H_i) = -1$ in_H_i **< 0 , 1** sinon

Et

$in\_O = 1*\omega_3 + H_1*\omega_4 + H_2*\omega_5$

O=1 si in_O **> 0**

O=-1 si in_O **< 0**

X_1	X_2	D
-1	-1	-1
-1	1	1
1	-1	1
1	1	-1

Pour classifier les exemples ci-dessus

Il suffit de prendre :

$\omega_3 = -1$

$\omega_4 = -1$

$\omega_5 = -1$

Vérification :

Ex1: (1,-1,-1); D = -1

$in\_H_1 = -3$;

$H_1 = f(-3) = -1$

$in\_H_2 = 1$;

$H_2 = f(1) = 1$

in_O = -1 +1 - 1 = -2

donc O= -1 qui correspond à D

Ex4: (1,1,1); D = -1

$in\_H_1 = 1$;

$H_1 = f(1) = 1$

$in\_H_2 = -3$;

$H_2 = f(-3) = -1$

in_O = -1 - 1 + 1 = -2

donc O= -1 qui correspond à D

Ex2: (1,-1,1); D = -1

$in\_H_1 = -1$;

$H_1 = f(-1) = -1$

$in\_H_2 = -1$;

$H_2 = f(-1) = -1$
$in\_O = -1 + 1 + 1 = 1$
donc O= +1 qui correspond à D

Ex3: (1,1,-1); D = -1
$in\_H_1 = -1$;
$H_1 = f(-1) = -1$
$in\_H_2 =-1$;
$H_2 = f(-1) = -1$
$in\_O = -1 + 1 + 1 = 2$
donc O = +1 qui correspond à D