

Examen final IN54 – Durée 2 heures **(Documents non autorisés)**

*Les 2 parties sont à rédiger sur des copies séparées*

**Partie 1 (14 points)**

**Exercice 1 (10 points)**

On souhaite implémenter un réseau de neurones multicouches pour résoudre un problème de classification de formes en deux classes. Chaque forme est caractérisée par deux attributs (X et Y). Le réseau de neurones répond 0 (respectivement 1) pour la classe C1 (respectivement C2).

1- Donner l'architecture du réseau de neurones, en précisant le nombre de neurones des couches d'entrée et de sortie, sachant qu'une seule couche cachée est utilisée. **(3 points)**

2- Etablir les règles d'apprentissage (celle des poids synaptiques liés à la couche de sortie et celle des poids liés à la couche cachée), en utilisant la méthode du gradient appliquée sur l'erreur en sortie du réseau. La sortie des neurones du réseau est calculée avec une fonction sigmoïde. **(7 points)**

**Exercice 2 (4 points)**

Un réseau de neurones entièrement connectés peut-être considéré comme une mémoire auto-associative. Expliquez.

**Partie 2 (6 points)**

**Théorie (2 pts) :**

Questions :

1. Expliquer ce qu'est le *Manifold Untangling*.
2. Décrivez aussi précisément que possible l'architecture classique d'un réseau de convolution (tel que LeNet 5 ou AlexNet par exemple) appliqué à la reconnaissance de formes dans les images.

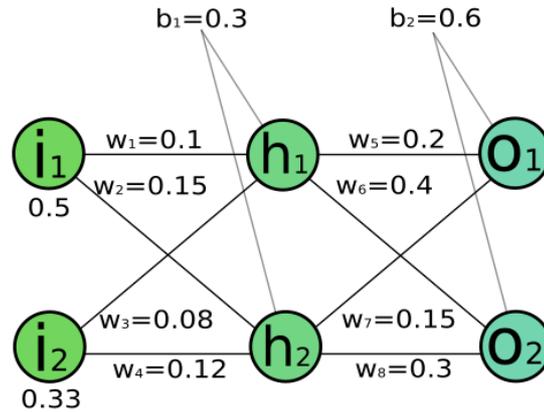
**Pratique (4 pts) :**

1. Soit la matrice d'entrée 7x7 suivante :

1	2	4	1	4	0	1
0	0	1	6	1	5	5
1	4	4	5	1	4	1
4	1	5	1	6	5	0
1	0	6	5	1	1	8
2	3	1	8	5	8	1
0	9	1	2	3	1	4

Quelle sera la sortie suite à un max pooling 3x3 avec un stride de 2 ?

2. Soit un réseau de neurones comprenant deux entrées, deux neurones cachés et deux neurones de sortie :



Les neurones cachés et de sortie comprennent un biais et ont une fonction d'activation de type logistique :

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

Calculer :

- La valeur en sortie des neurones  $O_1$  et  $O_2$
- L'erreur quadratique  $E_{O_1}$  et  $E_{O_2}$
- L'erreur totale  $E_{total}$

3. Supposons une entrée tridimensionnelle  $x = (x_1, x_2, x_3)$  connectée à 1 neurone avec comme fonction d'activation  $g_i$ . La propagation vers l'avant peut alors s'écrire :

$$z = \left( \sum_{k=1}^3 w_k x_k \right) + b$$

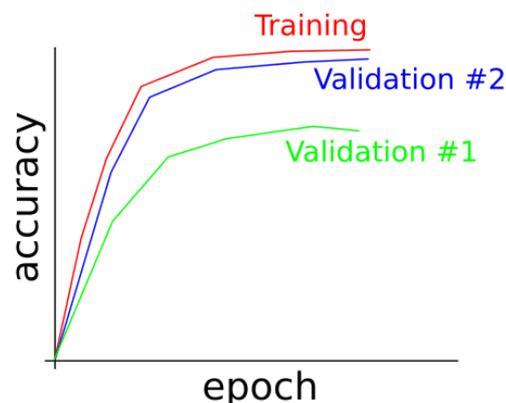
$$a_i = g_i(z)$$

Après entraînement de ce réseau, les valeurs des poids et du biais sont respectivement :  $w = (w_1, w_2, w_3) = (0.5, -0.2, 0.0)$  et  $b = 0.1$ .

Vous essayez 2 fonctions d'activation différentes  $g_1$  et  $g_2$  qui retournent respectivement les valeurs  $(a_1, a_2) = (0.67, 0.72)$ .

Quelles sont les 2 fonctions d'activation que vous avez testées (montrer par le calcul) ?

4. Les courbes ci-dessous illustrent la précision d'entraînement et de validation (x2) à chaque époque d'un apprentissage profond.



Quelle courbe indique un problème d'overfitting ? Donner au moins deux mesures permettant d'éviter l'overfitting lors d'un apprentissage.