2.6 Séance de TP sur la reconnaissance de chiffres

Remarques générales
Cette séance de travaux pratiques (TP) est l'occasion de mettre en œuvre les théories étudiées en cours. Il est donc indispensable d'avoir une bonne connaissance du cours pour analyser les résultats des expériences menées pendant la séance. Le sujet abordé pour ce troisième TP porte sur l'utilisation du filtre adapté et d'un réseau de neurones avec une couche cachée pour réaliser la discrimination dans un problème de reconnaissance des chiffres.

Les étudiants doivent construire leur raisonnement en s'appuyant sur une démarche expérimentale. Pour les aider dans cette démarche, ils sont encouragés à poser des questions aux enseignants pendant la séance. Ce TP donne lieu à un compte rendu écrit d'une page A4 (recto uniquement, taille des polices 11) qui est remis avant le jour suivant la séance de TP au format pdf.

Le rôle de ce compte rendu est de décrire les quelques points qui vous paraissent essentiels et non de faire une liste exhaustive des points abordés en TP. Chaque groupe travaille à son rythme et rédige son compte rendu en autonomie.

1/ **Réseau de neurones sur un exemple didactique dans $\mathbb{R}^2$**
On considère un problème à deux classes pour lequel les densités de probabilité (ddp) ressemblent à deux spirales dans $\mathbb{R}^2$. Comme pour ce premier exemple la dimension de $\boldsymbol{x}$ est $N = 2$, on peut représenter les vecteurs des bases d'apprentissage (voir figure 1 et 4) et des bases de généralisation (voir figure 2, 3 et 5) ainsi que les surfaces discriminantes des différentes techniques considérés (voir titre des figures 2, 3 et 5). Si besoin, l'annexe rappelle rapidement le fonctionnement d'un réseau de neurones.

Pour le réseau de neurones, figure 4 et 5, on représente en tireté les N_c droites décrites par le vecteur $\boldsymbol{w}_j$ (pour $j = 1, ..., N_c$). En effet, on peut afficher dans $\mathbb{R}^2$ la droite définie par :

$$\boldsymbol{w}_j^T \boldsymbol{x} = 0 \quad \forall j \tag{48}$$

Le vecteur $\boldsymbol{w}_j$ décrit les poids qui relient le vecteur $\boldsymbol{x}$ de la couche d'entrée au neurone de la couche cachée du réseau de neurones (voir diagramme représenté en annexe). Faites le lien entre ces droites et la forme de la surface discriminante du réseau de neurones. Commentez la forme des surfaces discriminantes des différents discriminateurs.

On considère une base de généralisation de taille P_{gen}. Commentez les performances des 3 discriminateurs considérés. Pour le réseau de neurones, expérimentez l'influence de N_c, le nombre de neurones dans la couche cachée, de η, le pas du gradient, et du nombre de passes [2]. Pour cette analyse, vous pourrez vous appuyer sur la figure 6 où est tracée l'évolution du critère J_T (voir Eq. (75)) en fonction des itérations du gradient stochastique ainsi que la figure 7 où est tracée l'évolution des taux d'apprentissage et de généralisation en fonction du nombre de passes du gradient stochastique.

Enfin, analysez l'influence de la taille de la base d'apprentissage. Validez votre analyse avec un encadrant et rédigez votre compte-rendu.

Formulez une ou deux questions pour votre compte-rendu.

2/ **Filtre adapté** (`question=2`)
On analyse un problème de discrimination à partir d'images de taille 7×6. Les éléments de la base de données suivent le modèle :

$$\boldsymbol{x} = \boldsymbol{\mu}_c + \boldsymbol{z} \quad \forall c \in \{0, 1, 2, ..., 9\} \tag{49}$$

où $\boldsymbol{\mu}_c$ est une image binaire de taille 7×6 décrivant un chiffre de 0 à 9 et $\boldsymbol{z}$ est une réalisation aléatoire d'un bruit blanc gaussien centré de variance 0.3. Quand $\boldsymbol{\mu}_c$ est connu, on peut mettre en œuvre le discriminateur fondé sur le filtre adapté :

$$\widehat{d}_{FA}(\boldsymbol{x}) = \arg\max_c (2\boldsymbol{\mu}_c^\dagger \boldsymbol{x} - \boldsymbol{\mu}_c^\dagger \boldsymbol{\mu}_c) \tag{50}$$

Analysez les performances de ce discriminateur en fonction du nombre de chiffres (voir ligne 65).

2. Le nombre de passes (epoch en Anglais) correspond au nombre de fois que le gradient a vu l'ensemble des vecteurs de la base d'apprentissage.

3/ **Comparaison avec le réseau de neurones** (`question=3`)
On utilise le réseau de neurones pour apprendre à reconnaître les chiffres à partir de $\boldsymbol{x}$. Commentez les performances du réseau de neurones en fonction du nombre de chiffres à retrouver (voir ligne 45). Validez votre analyse avec un encadrant et rédigez votre compte-rendu.

4/ **Filtre adapté avec amplitude inconnue** (`question=4`)
Le filtre adapté présenté ci-dessus suppose que l'amplitude des $\boldsymbol{\mu}_c$ est connue. Lorsque l'amplitude des $\boldsymbol{\mu}_c$ est inconnue, le modèle devient

$$\boldsymbol{x} = A\,\boldsymbol{\mu}_c + \boldsymbol{z} \quad \forall c \in \{0,\, 1,\, 2, ...,\, 9\} \tag{51}$$

où A est représente l'amplitude inconnue. Dans ce cas, le filtre adapté s'écrit

$$\widehat{d}_{FA_{ai}}(\boldsymbol{x}) = \arg\max_c \frac{(\boldsymbol{\mu}_c^\dagger \boldsymbol{x})^2}{\boldsymbol{\mu}_c^\dagger \boldsymbol{\mu}_c} \tag{52}$$

Comparez les performances des deux filtres adaptés dans le cas où les bases de généralisation sont générées suivant le modèle de l'équation (69). Observez-vous une différence significative ?

Ensuite, analysez la situation quand la base de généralisation est générée selon le modèle de l'équation (51) avec $A \sim \mathcal{U}[-5,\, 5]$ (enlever le commentaire ligne 71 du script). Validez avec un encadrant. Rédigez votre compte-rendu.

5/ **Réseau de neurones avec amplitude inconnue** (`question=5`)
Analysez les performances du réseau de neurones quand l'apprentissage est effectué avec une base qui vérifie le modèle de l'équation (69), mais que la base de généralisation est générée selon le modèle de l'équation (51) avec $A \sim \mathcal{U}[-5,\, 5]$. Validez avec un encadrant.

Analysez les performances quand la base d'apprentissage et la base de généralisation sont générées selon le modèle de l'équation (51) avec $A \sim \mathcal{U}[-5,\, 5]$ (enlever le commentaire ligne 50). Analysez l'influence du nombre de chiffres (ligne 45). Validez avec un encadrant.

Proposez une synthèse de ces résultats et rédigez votre compte-rendu. Cherchez une solution pour améliorer les performances du réseau de neurones.

Formulez une ou deux questions pour votre compte rendu.

Annexe : fonctionnement d'un réseau de neurones

Le schéma du réseau de neurones avec N_c neurones sur la couche cachée et N_s neurones sur la couche de sortie est représenté Figure 14. A l'entrée de chaque neurone $j \in \{1, ..., N_c\}$ de la couche cachée

$$\nu_j = \sum_{i=0}^{N} x_i w_{ji} = \boldsymbol{w}_j^T \boldsymbol{x} \tag{53}$$

où w_{ji} est le poids relatif à la connexion allant de l'entrée i vers le neurone j de la couche cachée.
A la sortie de chaque neurone j de la couche cachée on a :

$$y_j = f(\nu_j) = f\left(\boldsymbol{w}_j^T \boldsymbol{x}\right) \tag{54}$$

où la fonction d'activation f choisie pour ce TP est la sigmoïde définie par :

$$f(z) = \frac{1 - e^{-2z}}{1 + e^{-2z}} \tag{55}$$

A l'entrée de chaque neurone $k \in \{1, ..., N_s\}$ de la couche de sortie, on a :

$$\eta_k = \sum_{j=0}^{N_c} y_j w_{kj} = \boldsymbol{w}_k^T \boldsymbol{y} \tag{56}$$

où w_{kj} est le poids relatif à la connexion allant du neurone de la couche cachée j vers le neurone de sortie k. A la sortie de chaque neurone k de la couche de sortie, on a :

$$z_k = f(\eta_k) = f\left(\boldsymbol{w}_k^T \boldsymbol{y}\right) = f\left(\sum_{j=1}^{N_c} w_{kj} f\left(\sum_{i=1}^{N} w_{ji} x_i + w_{j0}\right) + w_{k0}\right) \tag{57}$$

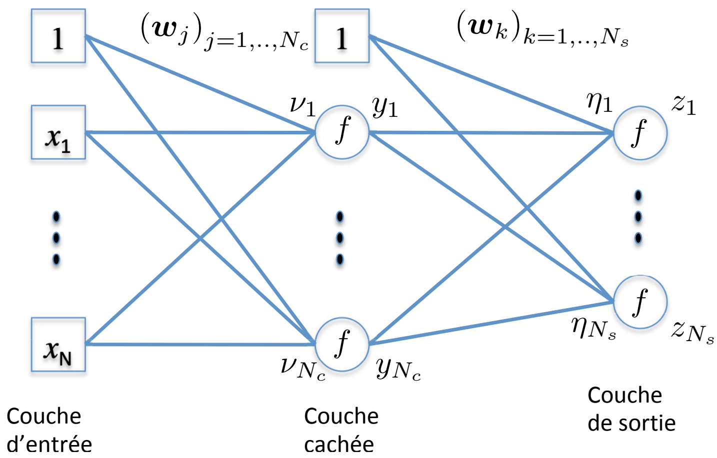

FIGURE 14 – Structure du réseau de neurones avec une couche cachée. Sur la couche d'entrée un vecteur de caractéristiques de taille N plus 1 offset, sur la couche cachée N_c neurones plus 1 offset et à la sortie N_s neurones.

Les paramètres $(\boldsymbol{w}_j)_{j=1..,N_c} = (w_{ji})_{j\in\{1,...,N_c\},i\in\{0,...,N\}}$ et $(\boldsymbol{w}_k)_{j=1..,N_s} = (w_{kj})_{k\in\{1,...,N_s\},j\in\{0,...,N_c\}}$ sont estimés avec une base d'apprentissage $\mathcal{B}_{app} = \left\{\left(\boldsymbol{x}_{app}^{(p)}, d_{app}^{(p)}\right); p \in \{1, ..., P_{app}\}\right\}$.

Le critère utilisé pour estimer les paramètres $\boldsymbol{w}_j$ et $\boldsymbol{w}_k$ est :

$$J_T = \frac{1}{P_{app}} \sum_{\boldsymbol{x}\in\mathcal{B}_{app}} J(\boldsymbol{x}, \boldsymbol{w}_j, \boldsymbol{w}_k) \quad \text{où} \quad J(\boldsymbol{x}, \boldsymbol{w}_j, \boldsymbol{w}_k) = \frac{1}{2}\sum_{k=1}^{N_s}(t_k - z_k)^2 \tag{58}$$

où $(t_k)_{k=\{1,..,N_s\}}$ est le vecteur des valeurs cibles des neurones de la couche de sortie pour le vecteur $\boldsymbol{x}$ de la base d'apprentissage. Dans ce TP, on a choisi les conventions suivantes.

— Si le nombre de classes est 2, alors $N_s = 1$ avec $t_1 = 1$ si $\boldsymbol{x}$ appartient à la classe 1 et $t_1 = -1$ sinon.

— Si le nombre de classes est supérieur à 2, alors N_s est égal au nombre de classe et $t_k = 1$ si $\boldsymbol{x}$ appartient à la classe k et $t_k = -1$ sinon.

Une fois les matrices $\boldsymbol{w}_j$ et $\boldsymbol{w}_k$ estimées, le réseau de neurones effectue la discrimination avec les règles suivantes.

— Si le nombre de classes est 2, alors $\widehat{d}_{RN}(\boldsymbol{x}) = \begin{cases} 1 & \text{si } z_1 \geq 0 \\ 2 & \text{sinon} \end{cases}$

— Si le nombre de classes est supérieur à 2, alors $\widehat{d}_{RN}(\boldsymbol{x}) = \arg\max_{k=1,\,...,\,N_s} z_k$

L'algorithme récursif utilisé pour estimer les paramètres $\boldsymbol{w}_j$ et $\boldsymbol{w}_k$ qui minimisent le critère J_T (voir Eq. (75)) est fondé sur un gradient stochastique. Celui-ci consiste à calculer le gradient de J pour chaque élément de la base d'apprentissage de manière itérative. Pour optimiser cet algorithme, il est souvent utile d'effectuer plusieurs "passes" à travers la base d'apprentissage.