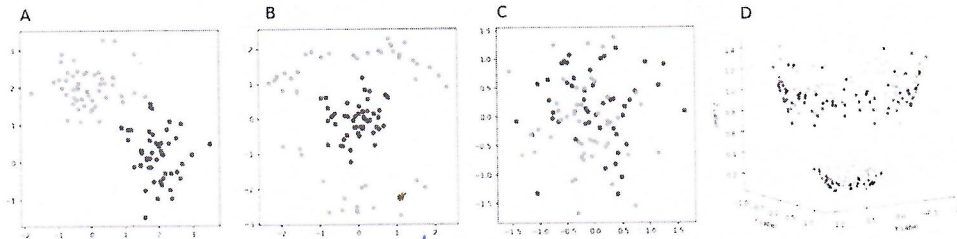


Réseaux de neurones fully-connected et convolutionnel

Règle QCM : Il y a une ou plusieurs bonnes réponses. Une bonne réponse vaut +1, une mauvaise réponse -1, une bonne réponse non choisie 0. Pas de pénalité pour les autres questions.

Nom : _____

1. Lesquels de ces plots de données sont *linéairement* séparables ? Entourez vos choix.



A. B. C. D. C, en l'état, n'est pas L.S.
Requiert "kernel trick"

2. Soit un perceptron qui reçoit quatre entrées $x = [1, 2, 3, 4]$ et dont la matrice de poids est $\omega = \begin{bmatrix} -0.5 \\ 0.5 \\ 0 \\ 1 \end{bmatrix}$.

Sachant que ce perceptron n'a pas de fonction d'activation et de biais, que vaut sa sortie ?

$1 \times 0.5 + 0.5 \times 2 + 0 + 4 = 4.5$ sortie d'un perceptron = un réel

3. Soit un réseau de neurones multi-classes conçu pour prendre des données $x \in \mathbb{R}^3$. Il est composé de deux couches cachées de deux perceptrons chacune. La dernière couche est connectée à 5 sorties ($y \in \mathbb{R}^5$).

(a) De combien de poids dispose le réseau (on omet les biais) ? Dessinez le réseau pour vous aider si nécessaire.

Taille de la couche $j = |j| \times |j-1|$. Il y a 3 couches de perceptron de taille 2, 2, 5 et une couche d'input de taille 3 : $3 \times 2 + 2 \times 2 + 2 \times 5 = 20$

(b) La sortie du réseau est passée dans la fonction softmax. Que vaut la somme de cette sortie ?

1

(c) La sortie 2 "softmaxée" vaut 0.70. Que représente cette valeur ?

- ☐ Le poids du perceptron responsable de cette sortie
☐ La loss pour la classe 2
☒ La confiance du réseau pour la classe 2
☐ On ne pourrait pas l'interpréter

* poids perceptron \neq sortie perceptron (même si corréli)
* la loss se calcule à partir de la sortie et d'une valeur attendue (cf exo 4)
* Si la sortie passe dans un softmax, elle représente la confiance ou probabilité sur 1 que la réponse soit correcte

4. Une loss de régression populaire s'appelle la "mean square error" (MSE). Elle se définit comme suit : $\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$ où \hat{y}_i est considérée comme la réponse attendue.

(a) Que vaut la MSE pour la réponse du réseau $y = [1, 2, 3]$ et $\hat{y} = [3, 1, 1]$

$$\frac{1}{3} (1-3)^2 + (2-1)^2 + (3-1)^2 = \frac{8}{3}$$

(b) Quels sont les rôles de l'algorithme de descente de gradient (backward pass) ?

- ☒ Minimiser la MSE
- ☐ Faire converger les poids des perceptrons vers 0
- ☒ Modifier les poids du réseau
- ☒ Que le réseau réponde $y = [3, 1, 1]$

Minimiser la loss = le réseau réponde $(3, 1, 1)$ (quand la loss va vers 0)

5. Soit un réseau convolutionnel (CNN) qui veut classifier des images RGB. Il est programmé pour prendre en entrée des images de taille $100 \times 100 \times 3$.

(a) Quelle est la dimension de l'entrée x du réseau (i.e. $x \in \mathbb{R}^?$) ?

- ☐ 100
- ☐ 300
- ☒ 30000

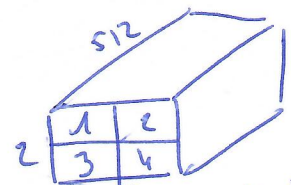
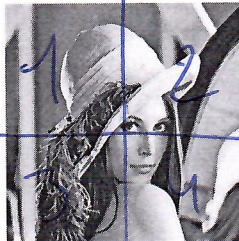
chaque composante RGB de chaque pixel est une entrée prise en compte par les filtres convolutifs

(b) Pour un glissement de 1, quelles tailles de filtres sont possibles pour la première couche de convolution ?

- ☒ $1 \times 1 \times 3$
- ☐ $99 \times 99 \times 1$
- ☒ $99 \times 99 \times 3$
- ☐ $1 \times 1 \times 1$

3^{es} dim. doit être égale à 3 : connectivité pleine en profondeur.

(c) L'image RGB ci-dessous représente Lenna, une mannequin du magazine Playboy dont la photo a été la première image utilisée par un algorithme de traitement d'image en 1973.



Le CNN a une connectivité spatiale locale (cf. gif vous)

Soit la sortie de la dernière couche de convolution de notre CNN de dimension $2 \times 2 \times 512$. Dessinez vous sur la photo originale la correspondance de ces 4 zones spatiales.

(d) Grâce à quelle propriété des CNN avez-vous répondu à la question précédente ?

- ☐ Connectivité pleine dans la profondeur
- ☐ Le max-pooling
- ☒ Connectivité spatiale locale

(e) Que représente la dimension de taille 512 dans la sortie mentionnée à la question c).

- ☒ L'encodage des caractéristiques pour chaque région spatiale
- ☐ Le nombre de poids d'un filtre de la couche de convolution précédente
- ☐ La taille du réseau fully-connected responsable de la classification à suivre

pour chaque zone de l'image, le CNN encode ses caractéristiques discriminantes nécessaires à la classification dans la 3^{es} dimension (dimension des "channels")