



Dr. Jehan-Antoine Vayssade

20 juillet 2025

Instructions

- **Durée** : 1 heure 30 min
- **Note totale / maximale** : 20 points
- **Les réponses doivent être argumentées**
- **Prenez uniquement un stylo**

1 Neurone de base (5 points)

1. (2 pts) Fournissez l'équation générale de calcul de sa sortie pour n entrées x_i , avec des poids w_i et un biais b . Introduisez la variable intermédiaire z et appliquez une fonction d'activation $f(z)$.
2. (3 pts) Considérez un neurone simple avec deux entrées $x_1 = 0.5$, $x_2 = -0.3$, des poids $w_1 = 0.8$, $w_2 = -1.2$, et un biais $b = 0.1$. Calculez manuellement :
 - (a) La sortie z et dite à quoi elle correspond ?
 - (b) La sortie $y = f(z)$ via la fonction d'activation relu.

Expliquez l'impact d'une valeur négative de z sur la sortie relu ainsi que son impact sur la rétropropagation.

2 Fonctions d'activation (3 points)

On considère les fonctions d'activation suivantes : Sigmoid, LReLU et ELU. Écrivez les dérivées de ces fonctions. Elles sont définies comme suit :

- **Leaky ReLU (LReLU)** avec un coefficient $\alpha \in (0, 1)$: $f(z) = \begin{cases} z & \text{si } z \geq 0 \\ \alpha z & \text{si } z < 0 \end{cases}$
- **Sigmoid** : $f(z) = \sigma(z) = \frac{1}{1+e^{-z}}$
- **ELU** : $f(z) = \begin{cases} z & \text{si } z \geq 0 \\ \alpha(e^z - 1) & \text{si } z < 0 \end{cases}$

Rappel : règles de dérivation élémentaires utiles

$$\frac{d}{dz} e^z = e^z, \quad \frac{d}{dz} \ln(1+u) = \frac{1}{1+u} \cdot u', \quad \frac{d}{dz} \frac{1}{z} = -\frac{1}{z^2}$$

3 Évaluation et optimisation (5 points)

1. (1 pt) Dans un dataset de 1000 échantillons, proposez une répartition raisonnable entre train, validation et test.
2. (1 pt) Cette répartition est-elle indispensable ? Quels sont les risques, bénéfices, contraintes et coûts associés ?
3. (1 pt) Sachant que, dans notre dataset, 10% des échantillons appartiennent à la classe 1, 20% à la classe 2 et 70% à la classe 3, que pouvez-vous en déduire sur le model appris ?
4. (2 pt) Pour le même cas, listez un ensemble de techniques ou de pratiques à mettre en œuvre.

4 Réseaux multicouches et apprentissage (4 points)

1. (2 pts) Illustrez avec un schéma simple (à dessiner à la main) un réseau de neurones à deux couches cachées. Indiquez les poids, biais et activations. Expliquez comment la non-linéarité des activations permet d'approximer des fonctions complexes.
2. (2 pts) Étendez le concept à plusieurs couches et déduisez en fonction de la "hauteur" et de la "largeur" du réseau, quels sont le degré et la précision de la fonction reconstruite.
3. (1 pt) Qu'est-ce que l'*overfitting* et l'*underfitting* ? Illustrez ces principes et proposez des techniques pour éviter ces problèmes.

5 Convolutions et architectures (3 points)

1. (1 pt) Dans un réseau convolutif (CNN), quel est l'avantage principal des convolutions par rapport à des couches fully connected ?
2. (1 pt) Expliquez l'effet de l'augmentation du nombre de filtres (kernels) dans une couche convolutionnelle sur :
 - (a) la capacité d'extraction de caractéristiques,
 - (b) le coût computationnel et la mémoire.
3. (1 pt) Appliquez manuellement une convolution 1D avec le kernel $K = [1, 0, -1]$ (détection de changements) sur le signal $S = [10, 20, 15, 30, 25]$, avec stride=1 et sans padding. Fournissez la sortie convolutive.