

# Examen IS5 - Introduction au Deep Learning

Polytech Lille

13 Decembre 2023

## Partie 1 : QCM

Pour chaque question suivante, veuillez entourer la réponse correspondant à votre choix. Il n'y a qu'UNE SEULE réponse correcte, sauf indication contraire. Aucune explication n'est nécessaire. Aucune pénalité ne sera appliquée pour une réponse incorrecte.

(1) Considérer un modèle de type Generative Adversarial Network (GAN) qui a été entraîné avec succès à générer des images de tortue. Laquelle des propositions suivantes est fausse ?

- (i) Le générateur vise à apprendre la distribution des images de tortue.
- (ii) Le discriminateur peut être utilisé pour classifier les images entre tortue et non-tortue.
- (iii) Après avoir entraîné le GAN, la fonction de coût associée au discriminateur atteint une valeur constante.
- (iv) Le générateur peut produire des images inédites de tortue.

(2) Parmi les techniques suivantes, lesquelles peuvent être utilisées pour réduire l'overfitting ? (Entourez la ou les réponses appropriées)

- (i) Data augmentation
- (ii) Dropout
- (iii) Early stopping
- (iv) Augmenter le nombre d'epoch

(3) Quel type de réseau de neurones est couramment utilisé pour la classification d'images ?

- (i) Recurrent Neural Network (RNN)
- (ii) Convolutional Neural Network (CNN)
- (iii) Feedforward Neural Network (FNN)
- (iv) Long Short-Term Memory (LSTM) Network

(4) Quels types de réseaux de neurones sont couramment utilisés pour la classification de séquences ? (Entourez la ou les réponses appropriées)

- (i) Recurrent Neural Network (RNN)
- (ii) Convolutional Neural Network (CNN)

- (iii) Feedforward Neural Network (FNN)
- (iv) Long Short-Term Memory (LSTM) Network
- (v) Transformer Encodeur

(5) Nous effectuons une convolution entre une image de taille 28 X 28 et un filtre de taille 7 X 7 avec un pas (stride) de 1. Quelle sera la taille de la matrice convoluée (feature map) ?

- (i) 20X20
- (ii) 21X21
- (iii) 22X22
- (iv) 25X25

(6) Parmi les propositions suivantes, lesquelles sont vraies à propos d'une couche de convolution ? (Entourez la ou les réponses appropriées)

- (i) Le nombre de paramètre dépend de la profondeur du volume d'entrée.
- (ii) Le nombre de biais est égal au nombre de filtres.
- (iii) Le nombre total de paramètres dépend du pas (stride).
- (iv) Le nombre total de paramètres dépend du padding.

(7) Quel est le rôle de la fonction d'activation dans un réseau de neurones ?

- (i) Normaliser les données
- (ii) Introduire de la non-linéarité
- (iii) Réduire l'overfitting
- (iv) Ajuster le taux d'apprentissage

(8) Parmi les fonctions d'activation suivantes, laquelle peut conduire au problème de vanishing gradient ?

- (i) ReLU
- (ii) Tanh
- (iii) Leaky ReLU
- (iv) Aucune des réponses précédentes

(9) Laquelle des fonctions suivantes peut être utilisée comme fonction d'activation dans la couche de sortie si l'on souhaite prédire les probabilités de n classes ( $p_1, p_2, \dots, p_n$ ) de telle sorte que la somme de p sur l'ensemble des n soit égale à 1 ?

- (i) Softmax
- (ii) ReLU
- (iii) Sigmoid
- (iv) Tanh

- (10) La tâche d'analyse des sentiments comme vu en cours est une tâche de prédiction de type many-to-many ?
- (i) Vrai
  - (ii) Faux
  - (iii) Peut-être vrai ou faux
  - (iv) Dépend du type de modèle utilisé
- (11) Quel est le principal avantage de l'utilisation d'un réseau de neurones récurrent (RNN) par rapport à un réseau de neurone feedforward ?
- (i) Traite les données séquentielles
  - (ii) Présente un plus grand nombre de couches cachées
  - (iii) Nécessite moins de paramètres
  - (iv) Présente de meilleures performances pour la classification d'images
- (12) Quel est le composant au sein d'un autoencoder responsable de la réduction des données d'entrée en une représentation compressée ?
- (i) Encoder
  - (ii) Decoder
  - (iii) KL divergence
  - (iv) La fonction de coût
- (13) À quoi le terme "attention" fait-il référence au sein du modèle Transformer ?
- (i) Concentrations sur les parties les plus pertinentes de l'entrée pendant le traitement.
  - (ii) Ajustement du taux d'apprentissage du modèle.
  - (iii) Définitions du nombre de couches du réseau.
  - (iv) Calcule le gradient pendant la rétropropagation.
- (14) Quelle architecture a introduit le concept de "skip connections" pour atténuer le problème de "vanishing gradient" ?
- (i) AlexNet
  - (ii) VGG
  - (iii) ResNet
  - (iv) Inception
- (15) Quel est le principal avantage de l'utilisation du "transfer learning" ?
- (i) Réduit le besoin de données étiquetées
  - (ii) Augmente la complexité du modèle
  - (iii) Diminue le temps de formation

- (iv) Améliore la généralisation
- (16) Qu'est-ce qui distingue un autoencodeur variationnel (VAE) d'un autoencodeur traditionnel ?
- (i) les VAE n'ont pas de couche bottleneck
  - (ii) Les VAE apprennent une distribution probabiliste dans l'espace latent.
  - (iii) Les VAE n'impliquent pas de fonction de perte de reconstruction
  - (iv) Les VAE sont spécifiquement conçues pour la classification d'images
- (17) Quel problème un VAE résout-il par rapport à un autoencodeur traditionnel ?
- (i) Overfitting
  - (ii) Réduction de dimension
  - (iii) Absence de capacité générative
  - (iv) Reconstruction de l'image

## Partie 2 : Questions

Veillez rédiger des réponses concises.

- (1) Vous êtes en train de concevoir un système d'apprentissage profond pour repérer la fatigue chez les conducteurs de voitures. Il est crucial que votre modèle puisse détecter efficacement la fatigue pour prévenir les accidents. Parmi les mesures d'évaluation suivantes, laquelle est la plus pertinente : Précision, Exactitude (Accuracy), Rappel, Coût ? Expliquez votre choix.
- (2) Quels sont deux avantages des réseaux de neurones à convolution par rapport aux réseaux entièrement connectés dans le cadre des tâches visuelles ?
- (3) Vous résolvez la tâche de classification binaire qui consiste à classer des images en chat ou non chat. Vous concevez un CNN avec un seul neurone de sortie. La sortie de ce neurone est  $z$ . La sortie finale de votre réseau,  $\hat{y}$ , est donnée par :

$$\hat{y} = \sigma(\text{ReLU}(z))$$

Vous classez toutes les entrées avec une valeur finale  $\hat{y} \geq 0.5$  comme des images de chats. Quel problème allez vous rencontrer ?

## Partie 3 : Architectures CNN

Considérez le réseau de neurones de type CNN défini par les couches énumérées dans la colonne de gauche du tableau ci-dessous. Veuillez compléter les dimensions associées au volume de sortie ainsi que le nombre de paramètres pour chaque couche. Exprimez les dimensions sous le format (H, W, C), où H, W et C représentent respectivement la hauteur, la largeur et le nombre de canaux. Sauf indication contraire, supposez un padding de 1 et un stride de 1 lorsque cela est approprié.

Notation :

1. CONV $x$ -N désigne une couche de convolution avec N filtres de hauteur et de largeur égales à  $x$ .

2. POOL-n désigne une couche de max-pooling  $n \times n$  avec un stride de n et un padding de 0.
3. FLATTEN applatit ses entrées, identique à `torch.nn.flatten` / `tf.layers.flatten`
4. FC-N désigne une couche entièrement connectée avec N neurones.

Table 1: Partie 3

Couche	Dimension	Volume	Activation	Nombre de paramètres
Input		32 X 32 X 3		0
CONV3-8				
Leaky ReLU				
POOL-2				
CONV3-16				
Leaky ReLU				
POOL-2				
FLATTEN				
FC-10				