

Examen: n°480252 "Machine Learning"

Copie n°12989014

Référence de l'examen: 251210-116221-202204-480252

Exam effectué depuis l'adresse IP 194.36.25.55 en utilisant le login OTP alexy88 / 5812

Commencé le 2025-12-11 15:31:10 et terminé le 2025-12-11 15:59:25

Note : 20/20 (52 points / 53) - 53 questions

Barème: (Barème 1) Demi-point attribué si au moins la moitié des bonnes réponses sont cochées

(Examen ouvert entre le 2025-12-11 15:30:00 et 2025-12-11 17:00:00)

Légende

Correct

Devait être coché

Faux

1 "Dans le clustering, un cluster correspond à un groupe de données..."

- A "aléatoires"
- B "étiquetées manuellement"
- C "séparées par classe"
- D "similaires entre elles"

(1 point / 1) Question à choix multiple

2 "Un dendrogramme permet de visualiser..."

- A "les gradients de mise à jour"
- B "les centroïdes des clusters"
- C "les valeurs de probabilité"
- D "le processus de fusion entre clusters"

(1 point / 1) Question à choix multiple

3 "La PCA sert à..."

- A "améliorer la précision des labels"
- B "augmenter le nombre de caractéristiques"
- C "produire des clusters sphériques"
- D "réduire la dimension des données"

(1 point / 1) Question à choix multiple

4 "Le DBI (Davies-Bouldin Index) est..."

- A "un algorithme de clustering"
- B "une fonction de perte"
- C "un modèle supervisé"
- D "une mesure de qualité des clusters"

(1 point / 1) Question à choix multiple

5 "Une instance est..."

- A "un vecteur de caractéristiques"
- B "une métrique de distance"
- C "un regroupement d'instances"
- D "un algorithme de clustering"

(1 point / 1) Question à choix multiple

6 "K-means converge généralement vers..."

- A "un optimum local"
- B "une solution déterministe"
- C "une solution aléatoire"
- D "l'optimum global"

(1 point / 1) Question à choix multiple

7 "Le centroïde d'un cluster est..."

- A "l'instance la plus proche du centre"
- B "un vecteur de poids"
- C "l'instance la plus fréquente"
- D "le point moyen des instances du cluster"

(1 point / 1) Question à choix multiple

8 "La normalisation des données permet..."

- A "de regrouper les clusters automatiquement"
- B "d'augmenter la dimension des données"
- C "d'équilibrer les échelles entre caractéristiques"
- D "de supprimer les zéros"

(1 point / 1) Question à choix multiple

9 "Pour choisir K avec DBI, on prend..."

- A "le K le plus proche de 1"
- B "le K le plus grand"
- C "le K avec le plus petit DBI"
- D "le K avec la plus grande inertie"

(1 point / 1) Question à choix multiple

10 "L'inertie intra-cluster correspond à..."

- A "le nombre de clusters"
- B "la moyenne des distances au centroïde"
- C "le nombre de points dans le cluster"
- D "le rayon moyen du cluster"

(1 point / 1) Question à choix multiple

11 "Le clustering hiérarchique commence avec..."

- A "un seul cluster"
- B "un cluster vide"
- C "le nombre optimal de clusters"
- D "un cluster par instance"

(1 point / 1) Question à choix multiple

12 "La distance la plus couramment utilisée en clustering est..."

- A "la distance cosine"
- B "la distance de Mahalanobis"
- C "la distance manhattan"
- D "la distance euclidienne"

(1 point / 1) Question à choix multiple

13 "Un bon DBI doit être..."

- A "proche de 1"
- B "grand"
- C "faible"
- D "proche de l'inertie"

(1 point / 1) Question à choix multiple

14 Le clustering appartient au domaine de l'apprentissage...

- A "semi-supervisé"
- B "supervisé"
- C "non supervisé"
- D "par renforcement"

(1 point / 1) Question à choix multiple

15 "Plus l'inertie est faible, alors..."

- A "il y a trop de clusters"
- B "le clustering est mauvais"
- C "le dataset est bruité"
- D "le clustering est meilleur"

(1 point / 1) Question à choix multiple

16 "K-means nécessite comme paramètre..."

- A "le nombre de clusters K"
- B "le modèle de linkage"
- C "le dendrogramme"
- D "le gradient"

(1 point / 1) Question à choix multiple

17 "Un neurone calcule..."

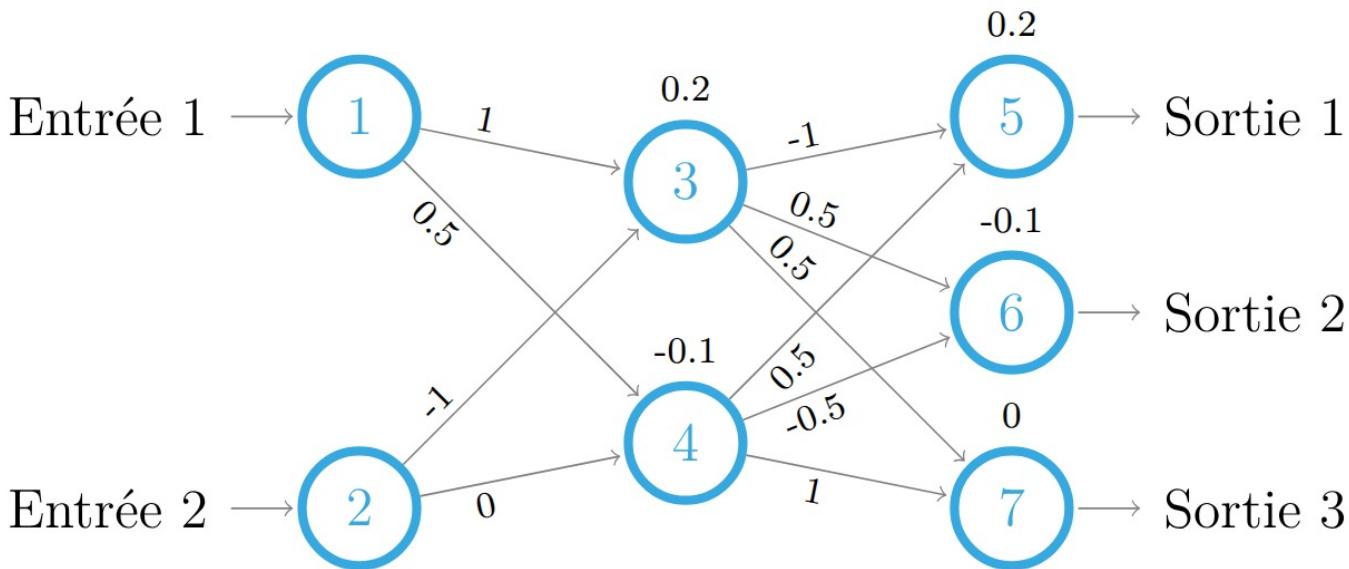
- A "un tirage aléatoire"
- B "une somme pondérée suivie d'une activation"
- C "une activation suivie d'une moyenne"
- D "une dérivée puis une somme"

(1 point / 1) Question à choix multiple

18 "La mse mesure..."

- A "la variance des données"
- B "le taux de fausses prédictions"
- C "la moyenne des carrés des erreurs"
- D "la moyenne absolue des erreurs"

(1 point / 1) Question à choix multiple



19 Voici un réseau de neurones avec une couche cachée de taille 2 et une couche de sortie de taille 3. Les poids sont donnés sur les arcs, et les biais au dessus des neurones. Chaque neurone est numéroté (entre 1 et 7).

Sachant que les entrées 1 et 2 sont respectivement fixées aux valeurs 0.4 et 0.6, et que la fonction d'activation est ReLu, déterminez la classe de sortie du réseau :

- A la classe 2 associée au neurone 7
- B la classe 1 associée au neurone 6
- C la classe 0 associée au neurone 5

(0 point / 1) Question à choix multiple

20 "Une fonction d'activation doit être..."

- A "aléatoire"
- B "constante"
- C "dérivable"
- D "linéaire"

(1 point / 1) Question à choix multiple

21 "La propagation avant consiste à..."

- A "calculer les sorties à partir des entrées"
- B "mettre à jour les poids"
- C "calculer l'erreur"
- D "mélanger les données"

(1 point / 1) Question à choix multiple

22 "Le réseau prédit la classe en..."

- A "calculant la moyenne"
- B "prenant l'index de la plus forte sortie"
- C "renvoyant un vecteur aléatoire"
- D "prenant la plus petite sortie"

(1 point / 1) Question à choix multiple

23 "La couche d'entrée d'un réseau de neurones fournit..."

- A "des étiquettes"
- B "des gradients"
- C "les valeurs des caractéristiques"
- D "des poids"

(1 point / 1) Question à choix multiple

24 "La descente de gradient ajuste les poids pour..."

- A "minimiser l'erreur"
- B "maximiser l'erreur"
- C "augmenter le nombre de neurones"
- D "minimiser les données"

(1 point / 1) Question à choix multiple

25 "L'apprentissage s'arrête généralement quand..."

- A "les sorties deviennent 1"
- B "l'erreur stagne"
- C "les poids deviennent nuls"
- D "le dataset change"

(1 point / 1) Question à choix multiple

26 "La fonction sigmoïde est définie par..."

- A "exp(x)"
- B "tanh(x)"
- C "max(0,x)"
- D "1 / (1 + exp(-x))"

(1 point / 1) Question à choix multiple

27 "Les poids d'un réseau déterminent..."

- A "la fonction d'activation"
- B "le nombre d'epochs"
- C "la dimension des données"
- D "l'influence des entrées sur les neurones"

(1 point / 1) Question à choix multiple

28 Voici les étapes, dans le désordre, de la procédure d'apprentissage d'un réseau de neurones pour la classification :

1. on ouvre une boucle itérant sur les instances ;
 2. on procède à la correction des poids et des biais ;
 3. on calcule l'erreur de sortie sur l'instance courante ;
 4. on ouvre une boucle itérant sur les époques ;
 5. on mélange l'ordre des instances ;
 6. on calcule la prédiction du réseau sur l'instance courante ;
 7. on procède à la rétropropagation de l'erreur dans le réseau ;
- Déterminez l'ordre des instructions.

- A 4, 5, 1, 6, 3, 7, 2
- B 4, 5, 1, 3, 6, 2, 7
- C 5, 1, 4, 6, 3, 7, 2
- D 1, 5, 4, 2, 3, 6, 7
- E 5, 1, 4, 6, 3, 2, 7

(1 point / 1) Question à choix multiple

29 "Un taux d'apprentissage trop grand..."

- A "assure la convergence"
- B "ne change rien"
- C "provoque une divergence"
- D "met les poids à zéro"

(1 point / 1) Question à choix multiple

30 "On considère un réseau dense avec 5 neurones en entrée, une couche cachée de 4 neurones, une seconde couche cachée de 3 neurones et 2 neurones en sortie. Chaque neurone hors couche d'entrée possède un biais. Combien de paramètres (poids + biais) possède le réseau ?"

- A "47"
- B "41"
- C "50"
- D "38"

(1 point / 1) Question à choix multiple

31 On veut évaluer les performances d'un modèle. Pour cela, on utilise la métrique de précision :

- A la précision est une fonction dérivable ;
- B la précision est une fonction exponentielle ;
- C le score est à minimiser : plus il est faible, mieux c'est ;
- D le score calculé est entre -1 et 1 ;
- E On calcule le nombre de bonnes réponses divisé par le nombre de réponses total ;

(1 point / 1) Question à choix multiple

32 On veut entraîner un modèle sur un problème. Il est nécessaire d'utiliser un ensemble de test, pourquoi ?

- A On l'utilise pour évaluer les performances du modèle après apprentissage.
- B On l'utilise pour supprimer la redondance dans les caractéristiques
- C On l'utilise pour arrêter l'apprentissage lorsque le modèle est suffisamment entraîné.
- D On l'utilise pour vérifier qu'il n'y a pas de données incohérente dans l'ensemble d'entraînement.

(1 point / 1) Question à choix multiple

33 "La couche de sortie, en classification, contient..."

- A "un neurone par classe"
- B "un neurone par image"
- C "un seul neurone"
- D "un neurone par caractéristique"

(1 point / 1) Question à choix multiple

34 "Un filtre appliqué à une image RGB est..."

- A "3D"
- B "1D"
- C "impossible"
- D "2D"

(1 point / 1) Question à choix multiple

35 "Avec un padding 'same', la sortie a..."

- A "une taille doublée"
- B "la même taille que l'entrée"
- C "une taille réduite de moitié"
- D "une taille aléatoire"

(1 point / 1) Question à choix multiple

36 "Le pooling sert à..."

- A "détecter des motifs"
- B "réduire la dimension spatiale"
- C "mélanger les données"
- D "augmenter la dimension"

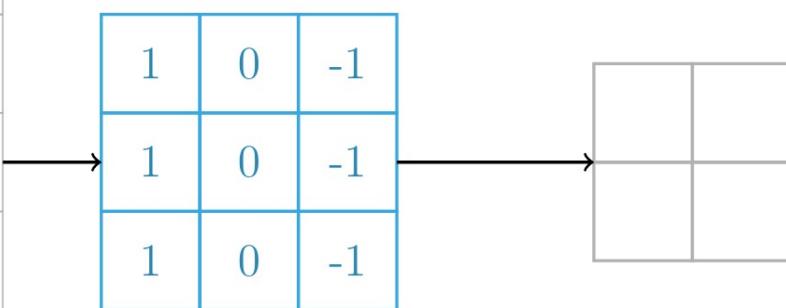
(1 point / 1) Question à choix multiple

37 "Le padding permet..."

- A "de normaliser l'image"
- B "de contrôler les dimensions de sortie"
- C "de supprimer les pixels"
- D "de supprimer le bruit"

(1 point / 1) Question à choix multiple

1	2	0	2	1
3	1	2	0	1
0	2	3	1	2
1	0	2	4	0
2	1	1	0	3

Image I (5x5)Noyau K (3x3)Sortie O (2x2)

38 On a un matrice d'entrée I de taille (5,5), et un noyau K de taille (3,3). Sachant que le stride est de 2, la matrice de sortie est :

- A [[-1, 1], [-3, 1]]
- B [[1, 1], [1, 1]]
- C [[2, 1], [-3, 0]]
- D [[1, 2], [0, 3]]

(1 point / 1) Question à choix multiple

39 "Le stride contrôle..."

- A "le nombre de canaux"
- B "le déplacement du noyau"
- C "le nombre de filtres"
- D "la fonction d'activation"

(1 point / 1) Question à choix multiple

40 "Le max pooling conserve..."

- A "la valeur maximale"
- B "la somme"
- C "la moyenne"
- D "la variance"

(1 point / 1) Question à choix multiple

41 "Une sortie 5x5x10 devient..."

- A "125 valeurs"
- B "25 valeurs"
- C "10 valeurs"
- D "250 valeurs"

(1 point / 1) Question à choix multiple

42 "Après convolution + pooling, on utilise..."

- A "une PCA"
- B "un réseau bayésien"
- C "une couche Flatten"
- D "une couche conv 1x1"

(1 point / 1) Question à choix multiple

43 "Une couche convolutionnelle avec M filtres produit..."

- A "3 images"
- B "une sortie aléatoire"
- C "M cartes de caractéristiques"
- D "1 image"

(1 point / 1) Question à choix multiple

44 "Le pooling n'a..."

- A "pas de paramètres"
- B "qu'un seul paramètre"
- C "un gradient compliqué"
- D "des milliers de paramètres"

(1 point / 1) Question à choix multiple

45 "Un CNN permet..."

- A "d'exploiter la structure spatiale des données"
- B "d'éviter les réseaux profonds"
- C "de supprimer la structure spatiale"
- D "d'aplatir automatiquement les données"

(1 point / 1) Question à choix multiple

46 "Une fenêtre de pooling 2x2 donne..."

- A "4 valeurs"
- B "2 valeurs"
- C "une valeur"
- D "0 valeur"

(1 point / 1) Question à choix multiple

47 "Un noyau de convolution est..."

- A "un vecteur de biais"
- B "un opérateur constant"
- C "une matrice de labels"
- D "une petite matrice de paramètres"

(1 point / 1) Question à choix multiple

48 "La couche Flatten transforme un tenseur en..."

- A "matrice 2D"
- B "image"
- C "vecteur 3D"
- D "vecteur 1D"

(1 point / 1) Question à choix multiple

49 "Augmenter le stride..."

- A "réduit la taille de sortie"
- B "duplique les valeurs"
- C "augmente la taille de sortie"
- D "ne change rien"

(1 point / 1) Question à choix multiple

50 "Dans un projet PyTorch, on définit une classe qui hérite de torch.utils.data.Dataset et on l'utilise avec un DataLoader. Quel est l'intérêt principal de cette approche ?"

- A "Charger les données à la volée en petits batches, sans tout garder en mémoire"
- B "Remplacer complètement le besoin d'un modèle nn.Module"
- C "Améliorer automatiquement la précision du modèle"
- D "Permettre d'entraîner sans définir de fonction d'erreur"

(1 point / 1) Question à choix multiple

51 "Dans un script PyTorch, on envoie le modèle et les tenseurs sur un 'device' (CPU ou GPU) avec la méthode .to(device). Pourquoi fait-on cela ?"

- A "Pour compresser automatiquement le modèle"
- B "Pour obliger PyTorch à utiliser uniquement la RAM"
- C "Pour empêcher la rétropropagation du gradient"
- D "Pour s'assurer que les calculs se font tous sur le même matériel (CPU ou GPU) et éviter les erreurs"

(1 point / 1) Question à choix multiple

52 "Dans la boucle d'entraînement PyTorch, on appelle successivement optimizer.zero_grad(), loss.backward() puis optimizer.step(). Que permet cette séquence d'instructions ?"

- A "Effacer les anciens gradients, calculer les nouveaux gradients, puis mettre à jour les poids du modèle"
B "Initialiser les poids à zéro, sauvegarder le modèle, puis l'envoyer sur le GPU"
C "Calculer la précision, afficher la loss, puis arrêter l'entraînement"
D "Normaliser les données, mélanger le dataset, puis recréer le DataLoader"

(1 point / 1) Question à choix multiple

53 "Dans un modèle PyTorch pour la classification d'images (comme MNIST), pourquoi enchaîne-t-on souvent des couches de convolution et de pooling, puis une couche Flatten suivie de couches denses ?"

- A "Les convolutions servent uniquement à réduire la taille, et la couche Flatten ajoute des paramètres au modèle"
B "Les couches denses sont inutiles, mais obligatoires pour utiliser un optimiseur"
C "Les convolutions extraient des motifs locaux, le pooling réduit la taille, la couche Flatten transforme les cartes de caractéristiques en vecteur pour les couches denses de classification"
D "Le Flatten sert uniquement à afficher les poids du modèle plus facilement"

(1 point / 1) Question à choix multiple