



Fleuret - Liebling
EE311 - SECTION
Spring 2020
Durée: 120 minutes



Archive Spring 2020

SCIPER: XXXXXX

Attendez le début de l'épreuve avant de tourner la page. Ce document est imprimé recto-verso, il contient 17 pages, les dernières pouvant être vides. Ne pas dégrafer.

- Posez votre carte d'étudiant sur la table.
- Les notes de cours ne sont pas autorisées.
- Seul est autorisé un "cheat-sheet" sous forme d'une feuille papier au format A4 avec des inscriptions (formulaire, résumés, etc.) au recto et au verso. Ce dernier doit impérativement avoir été préparé de manière personnelle. La technique d'élaboration est libre (manuscrit, dactylographié, collage, etc.) mais il ne pourra en aucun cas prendre la forme d'un assemblage d'artéfacts copié-collés directement du cours ou des labos (pas de micro-fiche, photocopies directes, etc.).
- L'utilisation d'une **calculatrice** et de tout outil électronique est interdite pendant l'épreuve.
- Les questions à choix multiples peuvent n'avoir que des déclarations incorrectes; dans ce cas, il faut impérativement remplir le carré "Aucune de ces réponses n'est correcte" ou la question sera considérée comme non répondue et pas de points ne seront ajoutés au score.
- La notation est telle que des réponses entrées aléatoirement résultent en un score nul:
 - Chaque question avec un simple choix parmi N réponses possibles donne +3pts lorsque la réponse est correcte, $-\frac{3}{N-1}$ pts si la réponse est incorrecte et 0pt s'il n'y a pas de réponse.
 - À l'intérieur de chaque question à choix multiples, la réponse à chacune des N cases donne +4/ N pt si l'état (coché ou vide) est correct et $-4/N$ pt s'il est incorrect. Si aucune case n'a été cochée la question est considérée comme non-répondue et donne 0pt.
- Une question pourra être annulée s'il s'avère par la suite qu'elle est incorrecte ou ambiguë.
- Utilisez un **stylo** à encre **noire ou bleu foncé** et effacez proprement avec du **correcteur blanc** si nécessaire.
- Si une question est erronée, l'enseignant se réserve le droit de l'annuler.

Respectez les consignes suivantes Observe this guidelines Beachten Sie bitte die unten stehenden Richtlinien		
choisir une réponse select an answer Antwort auswählen	ne PAS choisir une réponse NOT select an answer NICHT Antwort auswählen	Corriger une réponse Correct an answer Antwort korrigieren
ce qu'il ne faut PAS faire what should NOT be done was man NICHT tun sollte		

Première partie, questions à choix multiple

Pour chaque question marquer la case correspondante à la réponse correcte sans faire de ratures. Il n'y a qu'une seule réponse correcte par question.

Question 1 Le clustering est indiqué lorsque

- ☒ on dispose d'une large quantité de données non étiquetées
☐ on dispose d'une large quantité de données d'entraînement étiquetées

Question 2 Lequel des scénarios suivants conduit, en général, à une évaluation de la qualité d'un modèle (de classification) erronément élevé:

- ☒ évaluer le modèle entraîné sur le même jeu qui a servi à l'entraînement du modèle
☐ entraîner le modèle sur une seule classe avec un estimateur biaisé puis évaluer le modèle entraîné sur des données correspondant à une autre classe
☐ évaluer le modèle entraîné sur un jeu de données qui ne partage pas de données avec le jeu qui a servi à l'entraînement

Question 3 L'évaluation d'un test de dépistage produit les résultats suivants:

	Cancer	Pas de cancer	Total
Frottis +	190	210	400
Frottis -	10	3590	3600
Total	200	3800	4000

Quelle est la probabilité de ne pas avoir le cancer si le résultat du test est négatif?

- ☐ $200/4000 \approx 5\%$
☐ $3800/4000 \approx 95\%$
☒ $3590/3600 \approx 99.7\%$
☐ $10/190 \approx 5.3\%$

Question 4 Une validation croisée leave-one-out est:

- ☐ Une validation où l'on ignore la classe qui fournit les meilleurs résultats.
☒ Une validation croisée dont le nombre de folds est égal au nombre d'observations dans le jeu d'entraînement, et dont chaque fold est donc composé d'un jeu d'entraînement de taille $n - 1$ et d'un jeu de test de taille 1.
☐ Une validation d'un modèle où l'on croise tous les jeux de données qui ont servi à entraîner le modèle, sauf un.

Question 5 Lequel des aspects suivants représente un risque lors de la séparation d'un jeu de données en jeu d'entraînement et jeu de test?

- ☒ Que par malchance, seule une classe est représentée dans le jeu d'entraînement
☐ Que le jeu d'entraînement et le jeu de test n'aient pas d'éléments en commun
☐ Que le modèle apprenne à prédire toutes les classes avec une probabilité égale
☐ Que le temps d'entraînement soit excessivement long (jeu d'entraînement trop grand)

CORRECTION

Question 6 Un test de dépistage a pour but d'identifier si un échantillon (image, analyse sanguine) provient d'une personne atteinte d'une certaine maladie en le classifiant comme positif (malade) ou négatif (pas malade). Lorsque le test classe de manière incorrecte un échantillon d'une personne atteinte de la maladie comme étant en bonne santé, on appelle ce résultat un:

- ☐ vrai négatif
- ☐ faux positif
- ☒ faux négatif
- ☐ vrai positif

Question 7 Un algorithme de dépistage d'une maladie (positif correspond à malade) est ajusté tel que son rappel est égal à 1. Laquelle des affirmations suivante est correcte:

- ☒ le test ne produit aucun faux négatif
- ☐ un rappel (ou une sensibilité) de 1 indique que le test réduit efficacement le nombre de personnes auxquelles il faudra faire subir des examens approfondis
- ☐ tout test positif indique de manière fiable que la personne testée est vraiment malade

Question 8 Pour évaluer la qualité d'un modèle appris, on sépare communément les données en jeux de données tel qu'on ait:

- ☐ un jeu d'évaluation, un jeu de validation
- ☐ un jeu d'entraînement par classe
- ☐ pas de séparation du jeu de données (intégrité des données)
- ☒ un jeu d'entraînement, un jeu de validation, un jeu de test

Question 9 La matrice de confusion d'un problème de classification binaire

- ☐ un tableau avec autant de lignes et colonnes que de jeux d'entraînement distincts qui permet de comparer la performance du modèle évalué sur les données d'entraînement et de test.
- ☒ est un tableau avec 2 lignes et deux colonnes qui décompte les résultats selon qu'ils ont été classifié en TN, TP, FP, FN
- ☐ une matrice pour caractériser la moralité et la justice d'un modèle

Question 10 La sélection de variables pour des données à p variables consiste à

- ☒ éliminer un nombre $p - m$ de variables des données
- ☐ créer m nouvelles variables à partir des p variables dont on dispose initialement

Question 11 Après avoir entraîné une régression linéaire sur ses données Martine se rend compte que le modèle sous-apprend. Laquelle des actions suivantes serait adaptée?

- ☐ Simplifier le modèle en enlevant des variables et/ou utiliser un modèle polynomial avec un degré inférieur
- ☐ Garder le même modèle mais réduire la taille du jeu d'entraînement
- ☒ Rendre le modèle plus complexe en ajoutant des variables et/ou utiliser un modèle polynomial avec un degré plus élevé
- ☐ Aucune des actions proposées ci-dessus n'est adaptée au problème

CORRECTION

Question 12 Une agence de presse commissionne l'écriture d'un programme d'apprentissage automatique qui doit déterminer si un nouvel article entré dans la base de données d'articles traite ou non de sport sur la base de la fréquence des mots *ski*, *football*, *banque*, *police*, *salaire*, *politique*, *climat*, ou *éducation* contenus dans l'article. L'agence dispose de 5000 articles annotés avec les étiquettes *Suisse*, *monde*, *économie*, *divertissement*, *technologie*, *sport*, *sciences* et *santé*. Quel est le type de problème d'apprentissage automatique qui doit être résolu?

- ☐ apprentissage non-supervisé (réduction de dimension)
- ☐ apprentissage supervisé (régression linéaire)
- ☐ apprentissage non-supervisé (clustering)
- ☒ apprentissage supervisé (classification binaire)

Question 13 La fonction $k : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, $x, x' \mapsto k(x, x') = \max(x, x')$ est-elle un noyau?

- ☒ Non, car pour les observations $x^1 = 0$, $x^2 = 1$, et les valeurs $a_1 = -2$, $a_2 = 1$, on n'a pas : $\sum_{i=1}^2 \sum_{\ell=1}^2 a_i a_\ell k(x^i, x^\ell) \geq 0$.
- ☐ Oui, car pour les observations $x^1 = 0$, $x^2 = 1$, et les valeurs $a_1 = -2$, $a_2 = 1$, on a : $\sum_{i=1}^2 \sum_{\ell=1}^2 a_i a_\ell k(x^i, x^\ell) = -3$.

Question 14 Stéphanie a un problème de classification binaire. Ses données sont représentées en deux dimensions et elle considère l'ensemble de tous les cercles comme espace des hypothèses. A quelle expression cet espace correspond-il?

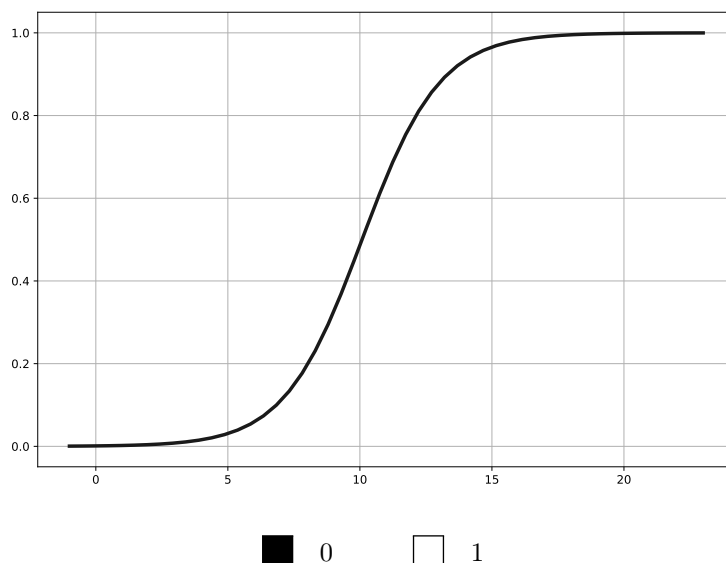
- ☐ $\mathcal{F} = \{\vec{x} \mapsto (x_1 - a)^2 + (x_2 - b)^2 - r^2\}$, où x_1 et x_2 sont les paramètres du modèle
- ☒ $\mathcal{F} = \{\vec{x} \mapsto (x_1 - a)^2 + (x_2 - b)^2 - r^2\}$, où a , b , et r sont les paramètres du modèle
- ☐ $\mathcal{F} = \{\vec{x} \mapsto ax_1 - x_2 + b\}$, où a et b sont les paramètres du modèle
- ☐ $\mathcal{F} = \{\vec{x} \mapsto \left(\frac{x_1}{a}\right)^2 + \left(\frac{x_2}{b}\right)^2 - 1\}$, où x_1 et x_2 sont les paramètres du modèle
- ☐ $\mathcal{F} = \{\vec{x} \mapsto \left(\frac{x_1}{a}\right)^2 + \left(\frac{x_2}{b}\right)^2 - 1\}$, où a et b sont les paramètres du modèle
- ☐ $\mathcal{F} = \{\vec{x} \mapsto ax_1 - x_2 + b\}$, où x_1 et x_2 sont les paramètres du modèle

Question 15 Laquelle des affirmations suivantes qui compare les avantages et inconvénients du choix d'une fonction de modélisation simple ou complexe est, en général, correcte?

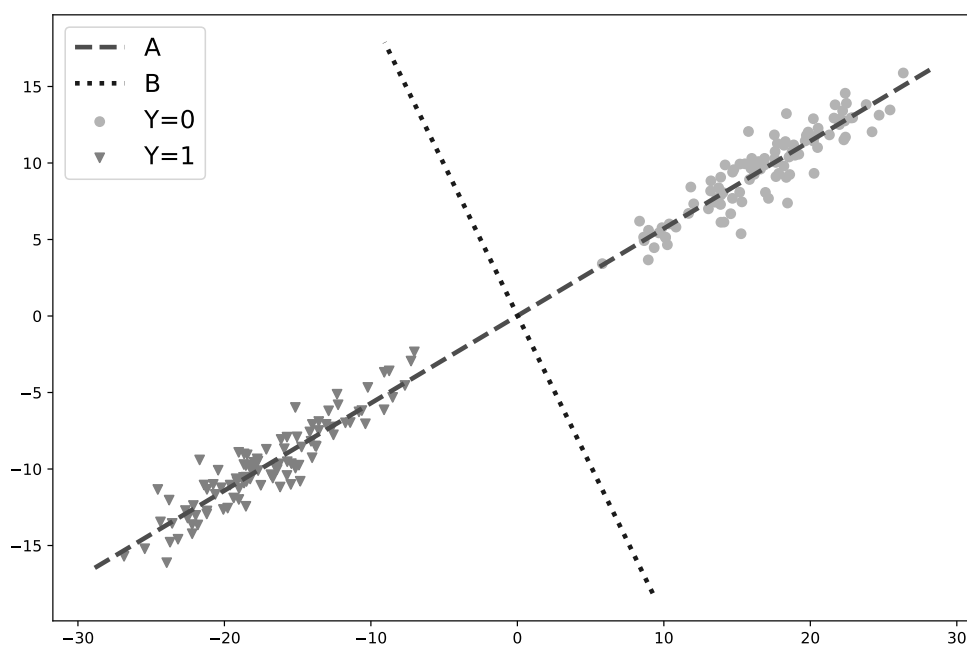
- ☒ un modèle simple permet d'éviter le sur-apprentissage, requiert moins de temps de calcul, permet une meilleure généralisation mais sa performance peut être moindre
- ☐ un modèle complexe permet d'éviter le sous-apprentissage, requiert plus de temps de calcul mais permet une meilleure généralisation et performance
- ☐ un modèle complexe permet d'éviter le sur-apprentissage, requiert plus de temps de calcul mais permet une meilleure généralisation et performance
- ☐ un modèle simple permet d'éviter le sous-apprentissage, requiert moins de temps de calcul mais sa capacité de généralisation et sa performance seront moindres

CORRECTION

Question 16 Soit un modèle de régression logistique entraîné sur un problème de classification binaire. La fonction logistique correspondant à la classe 1 est affichée ci-dessous. Quelle classe attribuera le modèle à une observation égale à 7?



Question 17 Dans la figure suivante, que représente la droite B?



☒ La frontière de décision d'un modèle de régression logistique

☐ La prédiction d'un modèle de régression logistique

☐ La prédiction d'un modèle de régression linéaire

Question 18 Soit un modèle $\vec{f} = X\vec{b}$ (X étant une matrice) associé à une fonction de coût $L = \left\| \vec{y} - \vec{f} \right\|_4^4$. Quelle ligne de code permet de calculer le gradient de la fonction de coût par rapport à \vec{b} ?

☐ `dL = 4 * X.dot(np.power(f-y, 3))`

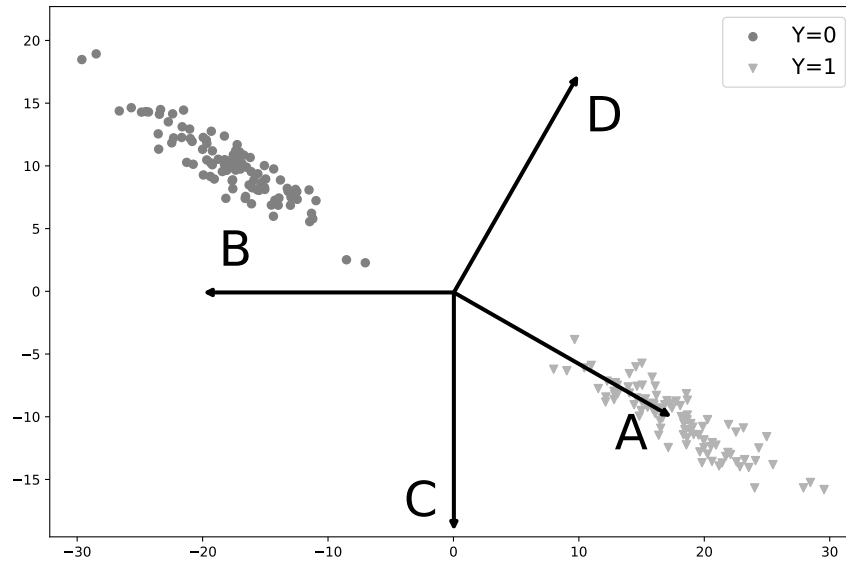
☒ `dL = 4 * np.power(f-y, 3) @ X`

☐ `dL = 4 * np.power(y-f, 3).dot(X)`

☐ `dL = -4 * X @ np.power(f-y, 3)`

CORRECTION

Question 19 Par ordre d'importance décroissante, quelles sont les composantes principales du jeu de données représenté ci-dessous?



☐ A, B
☐ A, C
☒ A, D

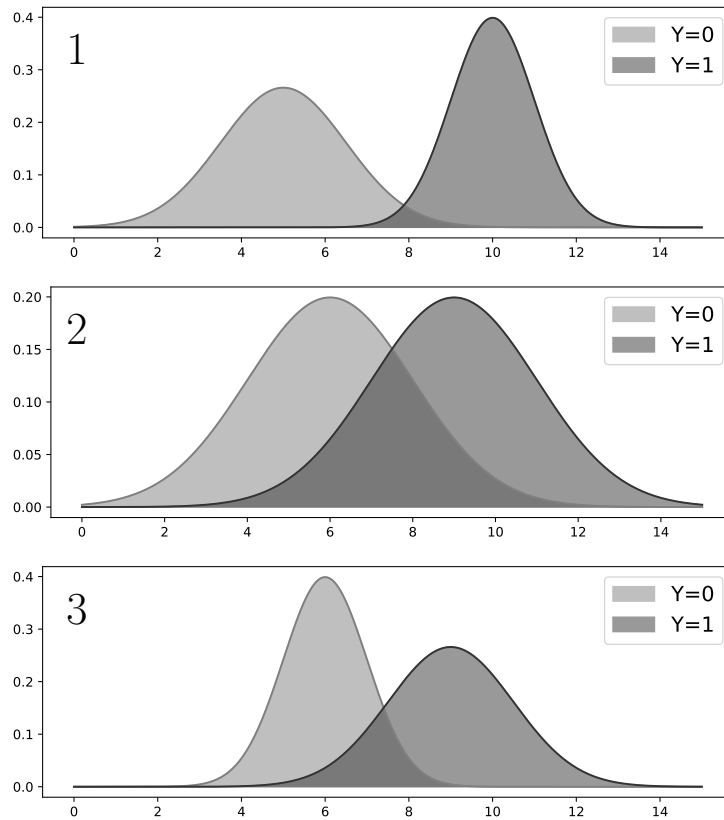
☐ B, A
☐ B, C
☐ B, D

☐ C, A
☐ C, B
☐ C, D

☐ D, A
☐ D, B
☐ D, C

CORRECTION

Question 20 Soit un jeu de données à 3 dimensions. Les densités de population des observations dans chacune des dimensions sont représentées ci-dessous. En appliquant une sélection de variables par filtrage avec le critère de corrélation, quel est l'ordre de pertinence des dimensions du jeu de données, par importance décroissante?



☐ 2, 3, 1

☐ 3, 1, 2

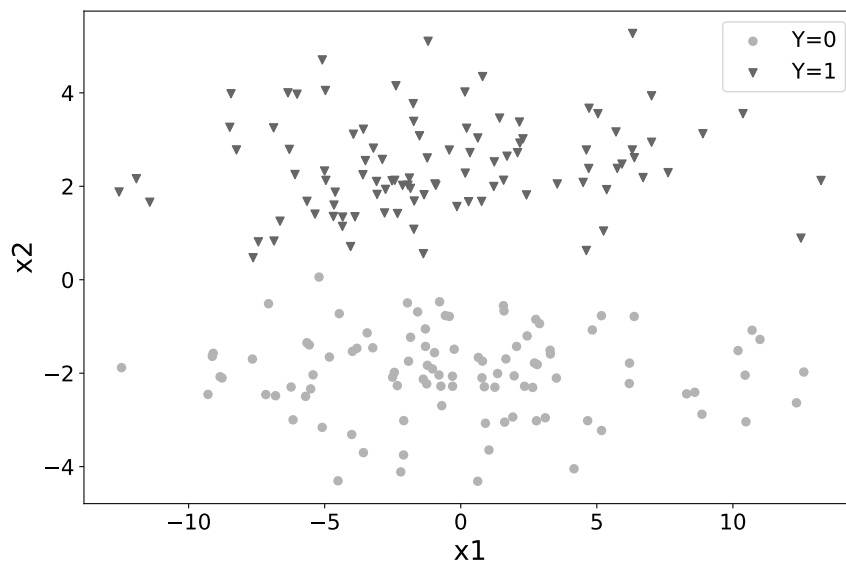
☐ 3, 2, 1

☐ 1, 2, 3

☐ 2, 1, 3

☒ 1, 3, 2

Question 21 Soit le jeu de données représenté ci-dessous. Quelle composante sera conservée si on le réduit à une dimension avec une méthode de filtrage utilisant le critère de corrélation?



☐ x_1

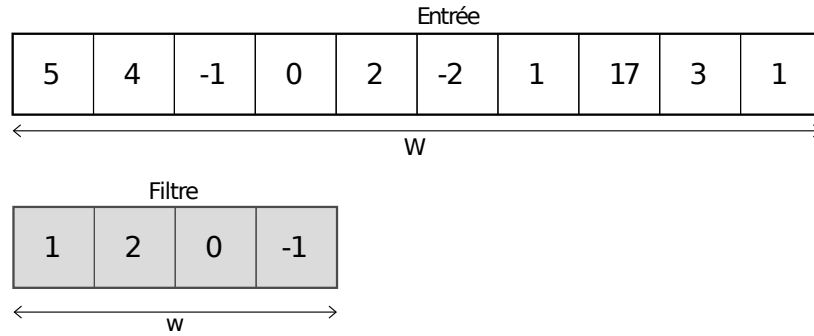
☒ x_2

CORRECTION

Question 22 Soit un réseau de neurones convolutif (CNN) dont l'entrée est une image RGB de taille $3 \times 250 \times 250$ et dont la sortie est un tenseur de taille 240×240 . Quel est le nombre de paramètres entraînables du kernel de cette première couche?

- ☐ 62'500 ☐ 243 ☒ 363 ☐ 300
☐ 121 ☐ 81 ☐ 100 ☐ 187'500

Question 23 En utilisant les conventions habituelles des réseaux de neurones convolutifs (CNN), quel est le résultat de la convolution suivante?



- ☐ [-5, -4, 13, 13, 0, 1, 3, -19, -3, 34, 23, 5, 1]
☒ [13, 0, 1, 3, -19, -3, 34]
☐ [-7, -2, 3, -3, 17, 39, 5]
☐ [5, 14, 7, -7, -2, 3, -4, 17, 39, 6, -15, -3, -1]

Question 24 Soit un réseau de neurones convolutif (CNN) dont l'entrée est une image RGB de taille $3 \times 250 \times 250$ et dont le kernel de la première couche est de taille $3 \times 15 \times 10$. Quelle est la taille de la sortie de cette première couche?

- ☐ $3 \times 236 \times 241$ ☒ 236×241 ☐ $3 \times 235 \times 240$
☐ $3 \times 250 \times 250$ ☐ 235×240 ☐ 250×250

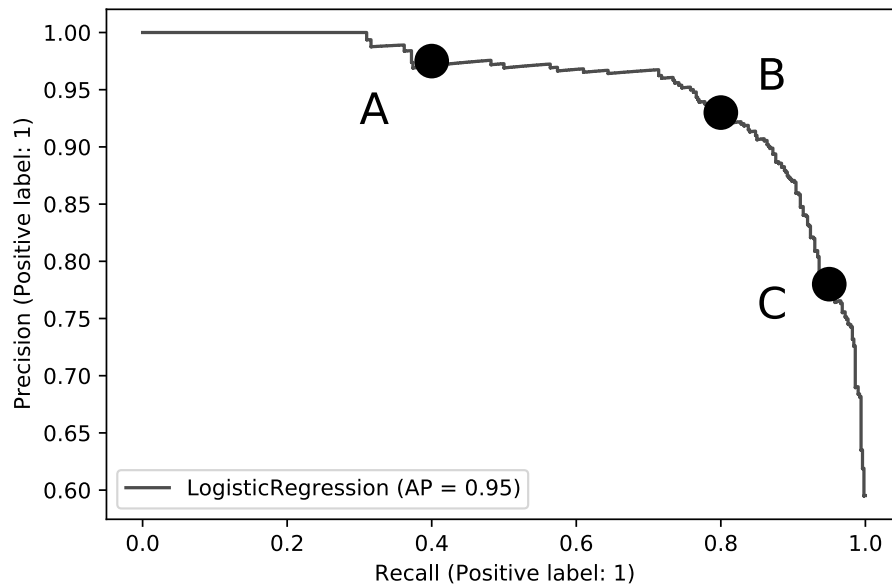
Question 25 Quelle est l'erreur empirique d'un modèle de classification ayant la matrice de confusion suivante?

14	3	2
9	23	4
3	9	33

- ☐ 1/4 ☐ 25 ☒ 0.3 ☐ 70%

CORRECTION

Question 26 Dans une voiture automatisée, un modèle d'apprentissage est utilisé pour détecter des piétons sur la rue (positif = piéton détecté). Selon la courbe de précision-rappel ci-dessous, quel modèle faut-il choisir?



☐ A ☐ B ☒ C

Question 27 Soit un classificateur binaire entraîné sur un jeu de données contenant 50 échantillons dans chaque classe. Sur base de sa matrice de confusion ci-dessous, quelle est son rappel? Le nombre de vrais positifs est 36.

36	14
4	46

☐ 92% ☒ 72% ☐ 82% ☐ 80% ☐ 90%

Question 28 Soit X un tableau numpy. Quelle ligne de code permet de calculer la somme de tous les éléments de X supérieurs à 10?

☐ `np.sum(X>10)`
☐ `np.sum(X[X>>10])`
☒ `X[X>10].sum()`

Question 29 Quelle ligne de code permet de connaître le nombre d'observations au sein du jeu de données X enregistré comme matrice numpy?

☒ `X.shape` ☐ `np.sum(X)`
☐ `X.size` ☐ `np.count(X)`

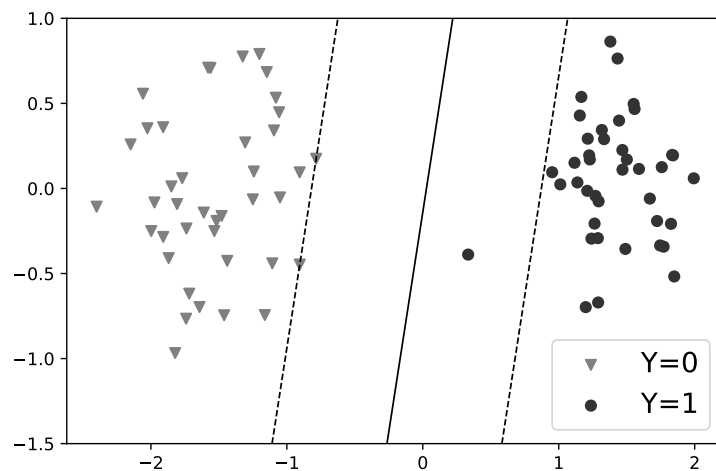
CORRECTION

Question 30 Quelle opération est effectuée dans le code suivant?

```
y = np.zeros(X.shape[1])
for i in range(X.shape[1]):
    for j in range(b.size):
        y[i] += b[j] * X[j, i]
```

☐ $y = X@b$ ☐ $y = X*b$ ☒ $y = b@X$ ☐ $y = b*X$

Question 31 Soit un modèle SVM entraîné sur un problème de classification binaire. Le jeu de données d'entraînement, l'hyperplan séparateur du modèle et ses marges sont représentés ci-dessous. Cette SVM est-elle à marge rigide?

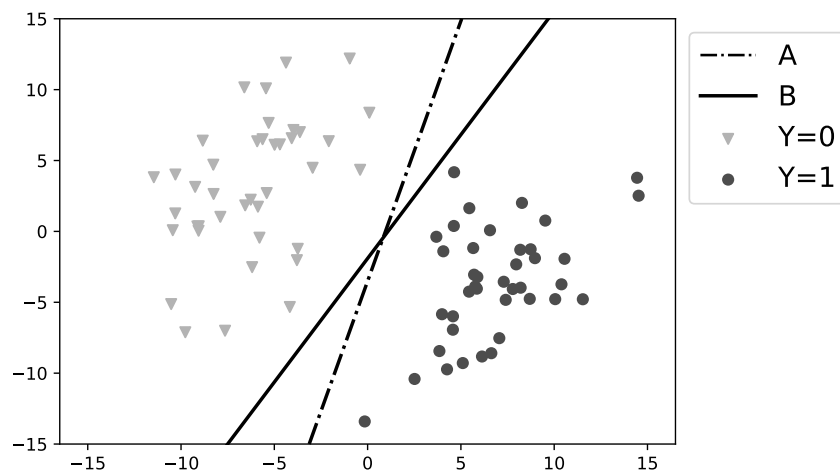
☐

Oui

☒

Non

Question 32 Laquelle des deux droites représente la frontière de décision d'une SVM?

☒

A

☐

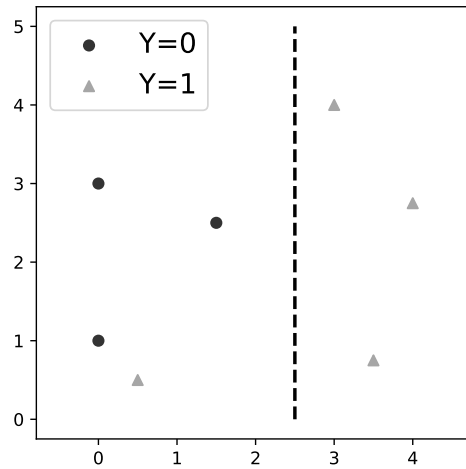
B

Question 33 Soit un modèle SVM de classification binaire défini par w le vecteur normal à son hyperplan séparateur et b son terme indépendant. Quelle ligne de code implémente la fonction de décision du modèle pour une entrée X ?

☒ $X.\text{dot}(w) > -b$ ☐ $X.\text{dot}(w) < b$ ☐ $X*w > b$

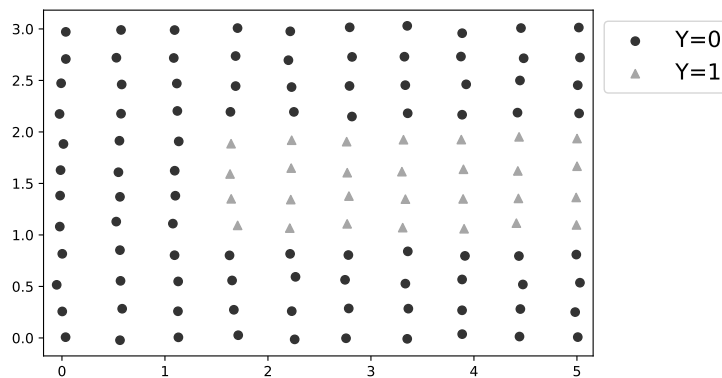
CORRECTION

Question 34 En employant l'impureté de Gini, quelle est la mesure d'impureté selon l'algorithme CART sur cette coupure?



☐ 0 ☐ 0.5 ☐ 1 ☒ 1.5

Question 35 Un arbre de décision est entraîné sur le jeu de données suivant, en utilisant l'algorithme CART. Quelle est la profondeur de l'arbre de décision ainsi créé ?



☐ 1 ☒ 3 ☐ 4 ☐ 6

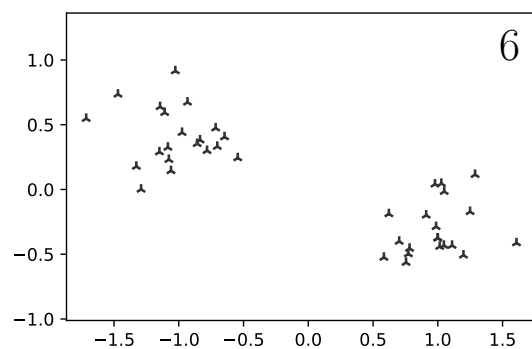
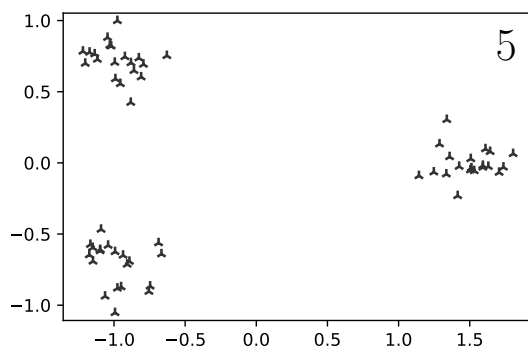
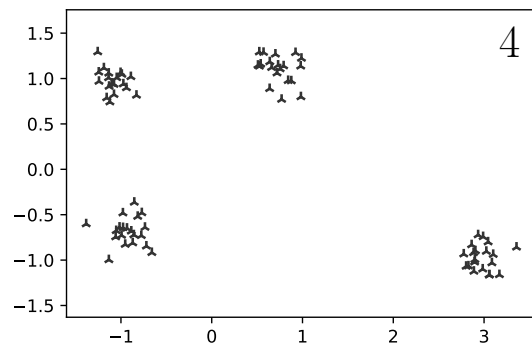
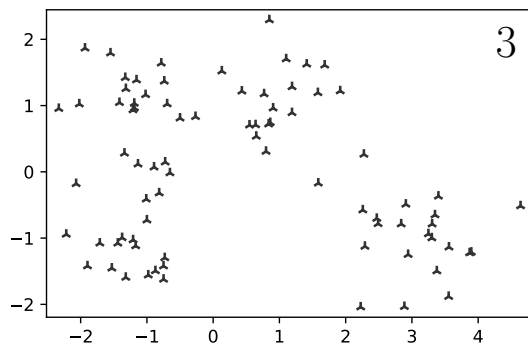
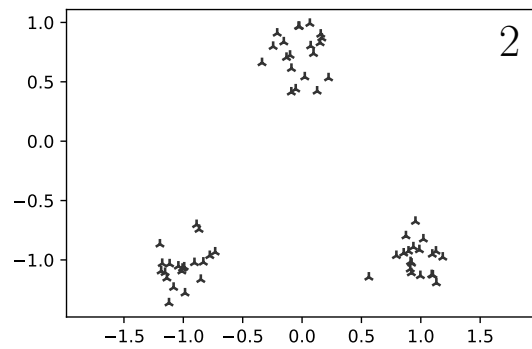
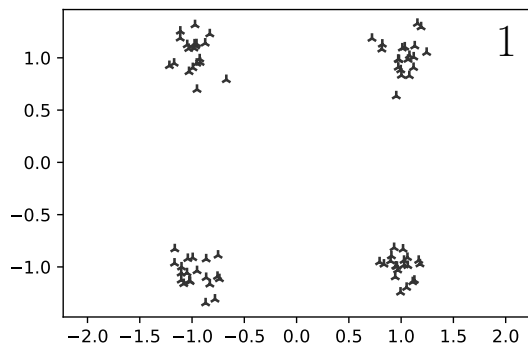
Question 36 Alice, Bob et Charlie sont en train de se balader, quand soudain ils aperçoivent un animal. Comme il fait sombre, ils n'arrivent pas bien à le distinguer. Le tableau suivant répertorie leurs impressions. En utilisant le bagging sur leurs prédictions, quel animal doit-on conclure qu'ils ont observé?

Animal	Alice	Bob	Charlie
Chien	98%	49%	22%
Chat	1%	0%	0%
Girafe	0%	0%	3%
Chamoix	1%	51%	75%

☐ Girafe ☒ Chamoix ☐ Chat ☐ Chien

CORRECTION

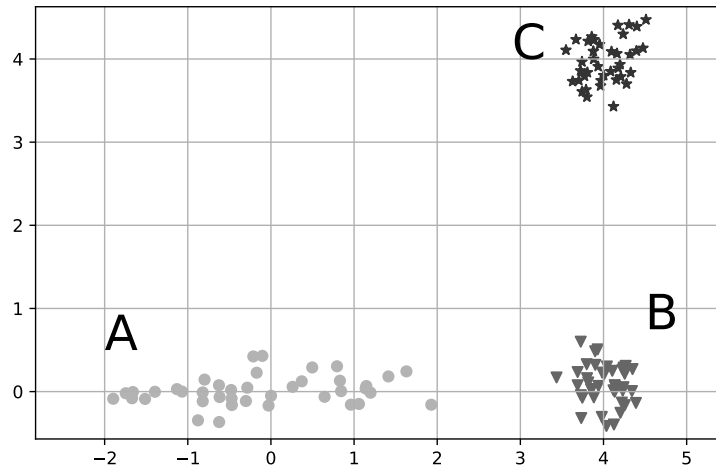
Question 37 Quel jeu de données correspond au dendrogramme suivant, construit en utilisant un clustering agglomératif à lien moyen?



☐ 1
 ☐ 2
 ☐ 3
 ☒ 4
 ☐ 5
 ☐ 6

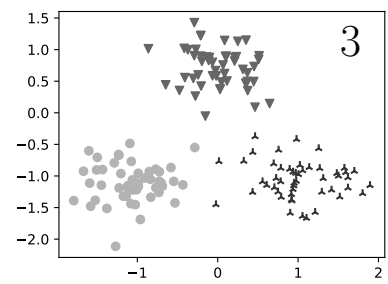
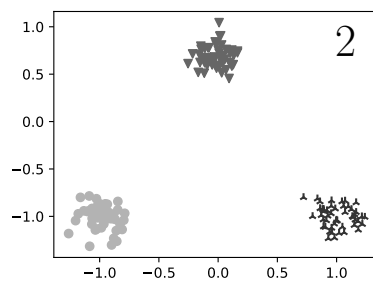
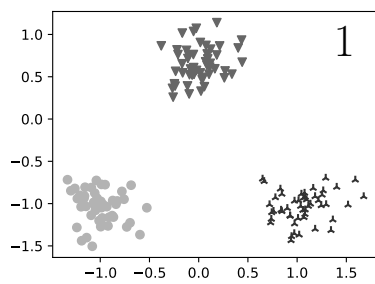
CORRECTION

Question 38 Soit le jeu de données séparé en trois clusters représenté ci-dessous. Selon un algorithme de clustering agglomératif avec lien simple, quels sont les deux prochains clusters à fusionner?



☐ A, C ☒ A, B ☐ B, C

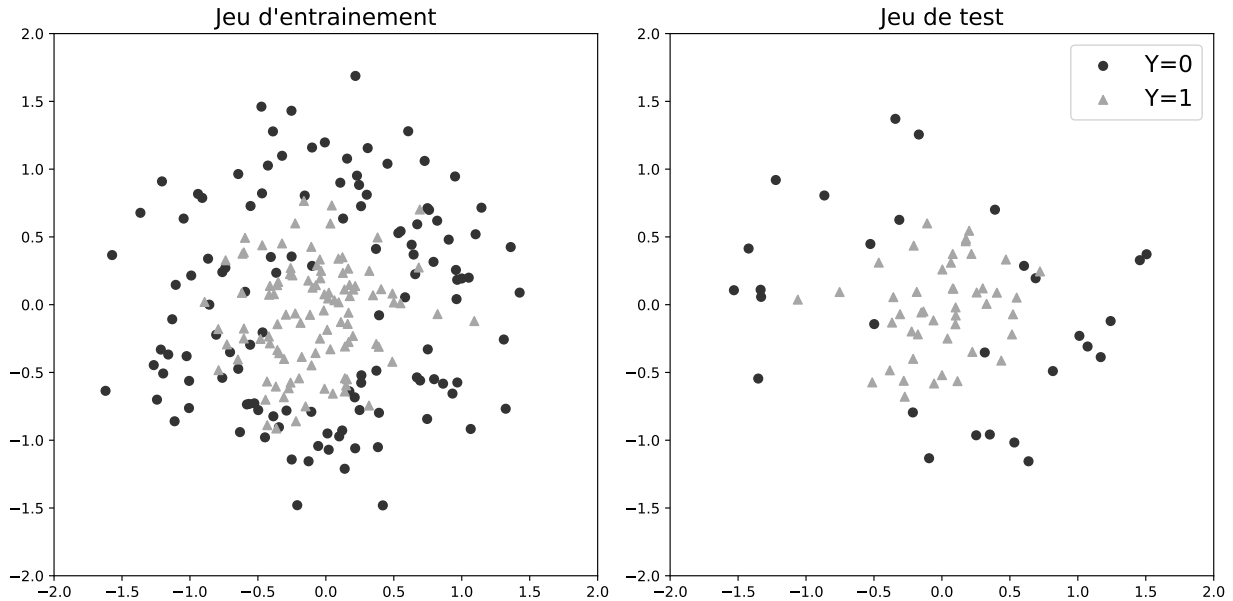
Question 39 Soient trois jeux de données séparés en trois clusters chacun, affichés ci-dessous. Trier les trois cas par coefficient de silhouette croissant.



☐ 2, 3, 1 ☐ 1, 3, 2 ☒ 3, 1, 2
☐ 1, 2, 3 ☐ 3, 2, 1 ☐ 2, 1, 3

CORRECTION

Question 40 Un réseau de neurones artificiel contenant suffisamment de paramètres et entraîné sur le jeu d'entraînement suivant peut-il obtenir une erreur nulle sur le jeu de test suivant?



☐ Oui ☒ Non

Question 41 En considérant la propagation vers l'avant du réseau de neurones artificiel suivant:

$$x^{(0)} \xrightarrow{W^{(1)}, b^{(1)}} s^{(1)} \xrightarrow{\sigma} x^{(1)} \xrightarrow{W^{(2)}, b^{(2)}} s^{(2)} \xrightarrow{\sigma} x^{(2)}$$

Avec $x^{(0)} \in \mathbb{R}^3$, $x^{(1)} \in \mathbb{R}^{100}$ et $x^{(2)} \in \mathbb{R}^2$; quel est le nombre de paramètres entraînables du modèle?

☐ 503 ☐ 500 ☒ 602 ☐ 603 ☐ 505 ☐ 105

Question 42 Dans la propagation du gradient vers l'arrière, les dérivées partielles par rapport à la sortie de la couche précédente k avant la fonction d'activation $\frac{\partial \ell}{\partial x_j^{(k)}}$ sont données par la fonction suivante:

$$\frac{\partial \ell}{\partial x_j^{(k)}} = \sum_i \frac{\partial \ell}{\partial s_i^{(k+1)}} w_{i,j}^{(k+1)}$$

Si les dérivées partielles $\frac{\partial \ell}{\partial s_i^{(k)}}$ sont regroupée dans tableau `d1_dskp1` et les poids $w_{i,j}^{(k+1)}$ dans une matrice `Wkp1`, quel code calcule le résultat de la formule précédente? (les résultats sont regroupés dans un tableau).

- ☐ `d1_dxk = d1_dskp1.dot(Wkp1.T)`
☒ `d1_dxk = Wkp1.T @ d1_dskp1`
☐ `d1_dxk = Wkp1.dot(d1_dskp1)`
☐ `d1_dxk = Wkp1 * d1_dskp1`

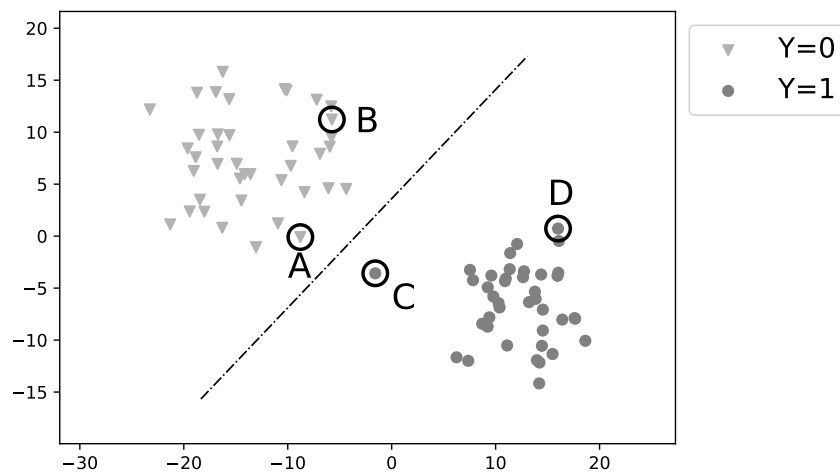
Deuxième partie, questions à choix multiples

Pour chaque question, marquer la case ou les cases qui correspondent à la réponse ou aux réponses correctes.

Question 43 ♣ Parmi les réponses suivantes, quelles sont des motivations pour réduire la dimension des données?

- ☒ réduire les coûts algorithmiques
- ☒ limiter le risque de sur-apprentissage
- ☒ exclure les variables qui ne sont pas pertinentes au problème et pourraient induire l'apprentissage en erreur
- ☒ limiter l'impact du "fléau de la dimension"
- ☒ visualiser les données
- ☐ Aucune de ces réponses n'est correcte.

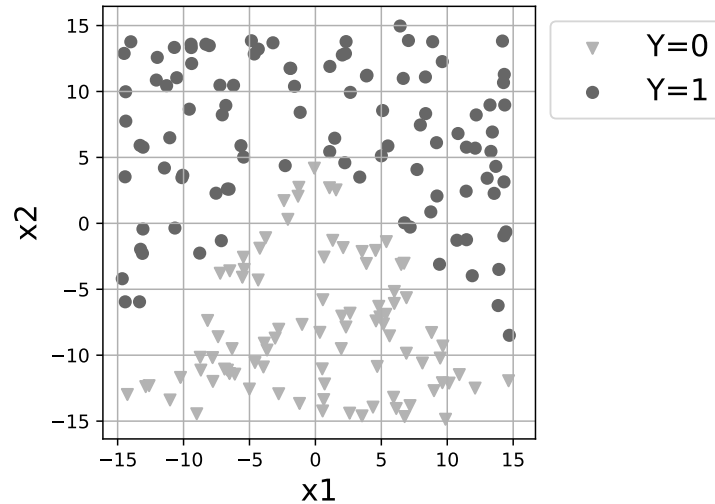
Question 44 ♣ Soit un modèle SVM entraîné sur un problème de classification binaire. Le jeu de données et l'hyperplan séparateur du modèle sont représentés ci-dessous. Parmi les points proposés, lesquels sont des vecteurs de support du modèle?



- ☒ A ☐ B ☒ C ☐ D ☐ Aucune de ces réponses n'est correcte.

CORRECTION

Question 45 ♣ Soit le jeu de données représenté ci-dessous. Quelle(s) transformation(s) de l'espace permet(tent) de le rendre linéairement séparable pour la classification binaire?


☒ $x_1, x_2 \rightarrow |x_1|, x_2$
☐ $x_1, x_2 \rightarrow x_1^2, x_2^2$
☐ $x_1, x_2 \rightarrow |x_1|, |x_2|$
☐ $x_1, x_2 \rightarrow x_1^2, x_2$
☐ $x_1, x_2 \rightarrow x_1, x_2^2$
☐ $x_1, x_2 \rightarrow |x_1|, x_2^2$
☐ $x_1, x_2 \rightarrow x_1, |x_2|$
☐ $x_1, x_2 \rightarrow x_1^2, |x_2|$
☐ Aucune de ces réponses n'est correcte.

Question 46 ♣ Quelle(s) ligne(s) de code permette(nt) de calculer correctement l'analyse en composantes principales (PCA) sur un jeu de données quelconque X?

☒ `values, vectors = np.linalg.eig(np.cov((X - X.mean(0)).T))`
☒ `values, vectors = np.linalg.eig(np.cov(X.T))`
☒ `_, s, Vt = np.linalg.svd(X - X.mean(0))`
☐ `_, s, Vt = np.linalg.svd(X)`
☐ Aucune de ces réponses n'est correcte.

Question 47 ♣ Soit X une matrice numpy contenant un jeu de données complet, et y le vecteur contenant les étiquettes correspondantes. Quelle(s) ligne(s) de code permette(nt) de calculer correctement le critère de corrélation utilisé pour une méthode de sélection de variables par filtrage?

☐ `np.abs((X - X.mean(0)).dot(y - y.mean()) / (X.std(0) * y.std() * y.size))`
☐ `((X - X.mean(0)) * (y - y.mean()) / (X.std(0) * y.std() * y.size))`
☐ `((X - X.mean(0)).T.dot(y - y.mean()) / (X.std(0) * y.std() * y.size))`
☒ `np.abs((X - X.mean(0)).T.dot(y - y.mean()) / (X.std(0) * y.std() * y.size))`
☐ Aucune de ces réponses n'est correcte.

CORRECTION

Question 48 ♣ La fonction XOR (ou exclusif) est une opération sur deux valeurs binaires X_1 et X_2 . Quand exactement une (et une seule) de ces deux valeurs est égale à 1 la fonction XOR renvoie 1, sinon elle renvoie 0. Considérons un simple réseau de neurones artificiel avec deux unités cachées et qui emploie ReLU comme fonction d'activation.

$$f(x; w, b) = \text{ReLU} \left(w^{(2)} \text{ReLU} \left(w^{(1)} x + b^{(1)} \right) + b^{(2)} \right) \quad (1)$$

Les poids $w^{(1)}$ et $b^{(2)}$ sont donnés par :

$$w^{(1)} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \text{ and } b^{(2)} = 0$$

Quels sont les poids $w^{(2)}$ et les biais $b^{(1)}$ qui font en sorte que le réseau de neurones artificiel final classifie correctement les données ?

☐ **a)** $b^{(1)} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ et $w^{(2)} = \begin{pmatrix} 0 & -1 \end{pmatrix}$

☒ **b)** $b^{(1)} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ et $w^{(2)} = \begin{pmatrix} 1 & -2 \end{pmatrix}$

☐ **c)** N'importe quel multiple scalaire de la réponse **a)**

☐ **d)** N'importe quel multiple scalaire de la réponse **b)**

☐ Aucune de ces réponses n'est correcte.

Question 49 ♣ Quelle(s) ligne(s) de code python calcule(nt) la solution optimale du problème de régression linéaire $\vec{y} = X\vec{b}$?

☐ `b = X.Ty np.linalg.inv(X.TX)`

☐ `b = X.T @ y @ np.linalg.inv(X.T @ X)`

☒ `b = np.linalg.inv(X.T @ X) @ X.T @ y`

☒ `b = np.linalg.inv(X.T.dot(X)).dot(X.T).dot(y)`

☐ Aucune de ces réponses n'est correcte.

Question 50 ♣ Soit un modèle de régression logistique appliqué à un problème de classification binaire. Si les observations du jeu de données d'entraînement possèdent 3 dimensions, quelle est la frontière de décision du classificateur?

☐ Un volume quelconque

☐ Un point

☐ Un volume convexe

☐ Une courbe non linéaire

☐ Une surface non linéaire

☒ Un plan

☐ Une droite

☐ Aucune de ces réponses n'est correcte.