Apprentissage et reconnaissance (GIF-4101 / GIF-7005)
Département de génie électrique et de génie informatique
Automne 2017

UNIVERSITÉ LAVAL

EXAMEN PARTIEL

Instructions : – Une feuille aide-mémoire recto verso manuscrite est permise ;
– Durée de l'examen : 2 h 50.
Pondération : Cet examen compte pour 35% de la note finale.

Question 1 (10 points sur 100)

Supposons que l'on fait du classement paramétrique selon deux classes et une variable en entrée (x scalaire), en modélisant les données de chaque classe par une loi normale :

$$p(x|C_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu_i)^2}{2\sigma^2}\right], i = 1{,}2.$$

La variance de chaque classe est la même, $\sigma_1 = \sigma_2 = \sigma$, alors que les moyennes μ_i et les probabilités a priori $P(C_i)$ sont différentes pour chaque classe. Sans perte de généralité, vous pouvez supposer que la moyenne de la classe 1 est inférieure à la classe 2, $\mu_1 < \mu_2$.
Avec cette modélisation, donnez l'équation analytique décrivant les frontières de décision entre les deux classes, en supposant une fonction de perte zéro-un (valeur égale pour les deux types d'erreurs). En une dimension, de telles frontières se résument à des seuils sur la valeur de x. Indiquez également de quelle façon le classement se fait dans les différentes régions séparées par les frontières.

Solution: L'équation du discriminant correspondant à la modélisation est :

$$\mathrm{h}_i(x) = \frac{p(x|C_i)\,P(C_i)}{p(x)}, i \in \{1{,}2\}.$$

La frontière de décision entre les deux classes correspond au point où $\mathrm{h}_1(x) = \mathrm{h}_2(x)$:

$$\begin{aligned}
\mathrm{h}_1(x) &= \mathrm{h}_2(x) \\
\frac{p(x|C_1)\,P(C_1)}{p(x)} &= \frac{p(x|C_2)\,P(C_2)}{p(x)} \\
p(x|C_1)\,P(C_1) &= p(x|C_2)\,P(C_2) \\
\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu_1)^2}{2\sigma^2}\right] P(C_1) &= \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu_2)^2}{2\sigma^2}\right] P(C_2) \\
\exp\left[-\frac{(x-\mu_1)^2}{2\sigma^2}\right] P(C_1) &= \exp\left[-\frac{(x-\mu_2)^2}{2\sigma^2}\right] P(C_2)
\end{aligned}$$

En appliquant le logarithme naturel de chaque côté de l'équation, on obtient :

$$\begin{aligned}
\ln\left[\exp\left[-\frac{(x-\mu_1)^2}{2\sigma^2}\right]P(C_1)\right] &= \ln\left[\exp\left[-\frac{(x-\mu_2)^2}{2\sigma^2}\right]P(C_2)\right] \\
\ln\exp\left[-\frac{(x-\mu_1)^2}{2\sigma^2}\right]+\ln(P(C_1)) &= \ln\exp\left[-\frac{(x-\mu_2)^2}{2\sigma^2}\right]+\ln(P(C_2)) \\
-\frac{(x-\mu_1)^2}{2\sigma^2}+\ln(P(C_1)) &= -\frac{(x-\mu_2)^2}{2\sigma^2}+\ln(P(C_2))
\end{aligned}$$

$$\begin{aligned}
\frac{1}{2\sigma^2}\left[(x-\mu_2)^2-(x-\mu_1)^2\right]+\ln(P(C_1))-\ln(P(C_2)) &= 0 \\
\frac{1}{2\sigma^2}\left[x^2-2x\mu_2+\mu_2^2-(x^2-2x\mu_1+\mu_1^2)\right]+\ln\left[\frac{P(C_1)}{P(C_2)}\right] &= 0 \\
\frac{1}{2\sigma^2}\left[-2(\mu_2-\mu_1)x+\mu_2^2-\mu_1^2\right]+\ln\left[\frac{P(C_1)}{P(C_2)}\right] &= 0 \\
\frac{1}{2\sigma^2}\left[-2(\mu_2-\mu_1)x+(\mu_2-\mu_1)(\mu_2+\mu_1)\right]+\ln\left[\frac{P(C_1)}{P(C_2)}\right] &= 0 \\
\frac{\mu_2-\mu_1}{2\sigma^2}\left[-2x+\mu_1+\mu_2\right]+\ln\left[\frac{P(C_1)}{P(C_2)}\right] &= 0
\end{aligned}$$

$$\begin{aligned}
-2x &= -\mu_1-\mu_2-\frac{2\sigma^2}{\mu_2-\mu_1}\ln\left[\frac{P(C_1)}{P(C_2)}\right] \\
x &= \frac{\mu_1+\mu_2}{2}+\frac{\sigma^2}{\mu_2-\mu_1}\ln\left[\frac{P(C_1)}{P(C_2)}\right]
\end{aligned}$$

La frontière de décision consiste donc en un seuil θ donné par :

$$\theta = \frac{\mu_1+\mu_2}{2}+\frac{\sigma^2}{\mu_2-\mu_1}\ln\left[\frac{P(C_1)}{P(C_2)}\right],$$

où les données pour lesquelles $x < \theta$ sont assignées à la classe C_1 et autrement à la classe C_2.

Question 2 (20 points sur 100)

Soit une loi de Bernoulli de paramètre p, soit la probabilité de succès (probabilité d'obtenir une valeur 1). Une loi géométrique de paramètre p modélise le nombre de répétitions de tirages selon une loi de Bernoulli de paramètre p nécessaire pour obtenir un premier succès. La loi s'exprime selon la probabilité $P(x|p)$ d'obtenir un premier succès après x tirages :

$$P(x|p) = (1-p)^{x-1}\,p.$$

La variable x est un entier positif ($x = 1,2,\ldots$). L'espérance d'une variable X suivant une loi géométrique de paramètre p est $\mathbb{E}(X) = \frac{1}{p}$ alors que sa variance est de $\mathrm{Var}(X) = \frac{1-p}{p^2}$.

(10) (a) Calculez la fonction pour estimer par un maximum de vraisemblance l'espérance d'une loi géométrique sur un jeu $\mathcal{X} = \{x^t\}_{t=1}^N$ de N échantillons. Donnez les développements analytiques complets pour en arriver à votre réponse.

Remarque : On traite le développement d'estimateur selon le maximum de vraisemblance avec une loi de probabilité de la même façon qu'on le fait avec une loi de densité de probabilité.

Solution:

$$
\begin{aligned}
L(\mathcal{X}|p) &= \log P(\mathcal{X}|p) = \log \prod_{t=1}^{N} P(x^t|p) \\
&= \sum_{t=1}^{N} \log \left[(1-p)^{x^t-1} p\right] \\
&= N \log p + \log(1-p) \sum_{t=1}^{N} (x^t - 1), \\
\frac{\partial L(\mathcal{X}|p)}{\partial p} &= N \frac{\partial}{\partial p} \log p + \left(\sum_{t=1}^{N} (x^t - 1)\right) \frac{\partial}{\partial p} \log(1-p) = 0 \\
&= \frac{N}{p} + \left(\sum_{t=1}^{N} x^t - N\right) \frac{1}{1-p}(-1) = 0, \\
\frac{N}{p} &= \frac{1}{1-p} \left(\sum_{t=1}^{N} x^t - N\right), \\
\frac{1-p}{p} &= \frac{1}{N} \left(\sum_{t=1}^{N} x^t - N\right), \\
\frac{1}{p} &= 1 - 1 + \frac{1}{N} \sum_{t=1}^{N} x^t, \\
\hat{p} &= \frac{1}{\frac{1}{N}\sum_{t=1}^{N} x^t}.
\end{aligned}
$$

(5) (b) Nous avons accès à K machines à sous distinctes dans un casino. Chaque machine est programmée pour fournir un montant m_i, connu des joueurs et différent pour chaque machine. Une partie avec une machine correspond à une expérience aléatoire où l'on fait le tirage d'une valeur binaire x_i suivant une loi de Bernoulli de paramètre p_i, qui est différent pour chaque machine et inconnu des joueurs. Une valeur de $x_i = 1$ signifie que le montant m_i est donné au joueur alors qu'autrement une valeur $x_i = 0$ est obtenue.

Supposons qu'une estimation $\hat{p}_i$, $i = 1,\ldots,K$, des paramètres de chaque machine est

disponible. Donnez l'équation permettant de choisir la machine qui maximise les gains espérés du joueur en justifiant votre réponse.

Solution: Maximiser le gain espéré du joueur à chaque partie correspond à maximiser l'espérance $\mathbb{E}(m_i\,\hat{p}_i)$. Comme le gain m_i et l'estimateur $\hat{p}_i$ sont constants, la décision optimale consiste à sélectionner la machine avec la valeur $m_i\,\hat{p}_i$ maximale :

$$x^* = \underset{i=1,\ldots,K}{\operatorname{argmax}}(m_i\,\hat{p}_i).$$

(5) (c) Supposons maintenant qu'un joueur n'a pas accès aux estimations $\hat{p}_i$ des paramètres des machines à sous. La stratégie est de supposer des paramètres initiaux des machines $\hat{p}_i$ arbitraires et identiques pour toutes les machines. Tant qu'aucun gain n'est fait, les machines sont jouées à tour de rôle. Lors du premier gain par une machine, l'estimation $\hat{p}_i$ de la machine gagnante est mise à jour à partir du gain observé et remplace le paramètre initial. Pour les gains subséquents d'une même machine, l'estimation $\hat{p}_i$ se fait à partir des gains historiques observés. De plus, la prise de décision se base sur la maximisation du gain espéré, une fois le premier gain est observé.

Cette approche pour la prise de décision comporte des problèmes importants. Expliquez dans votre cahier les principaux éléments problématiques. Proposez également des correctifs permettant de régler le problème.

Solution: Un problème avec cette approche est que l'on va prendre des décisions basées sur des estimations faites avec très peu d'observations, qui sont donc peu fiables. De plus, les décision passent au choix de l'action maximisant le gain espéré après une seule observation de gain, rendant la décision fortement dépendante de l'estimation des paramètres $\hat{p}_i$ des actions. Une mauvaise estimation de paramètres basés sur peu de données peu condamner une machine à prendre des décisions sous-optimales, comme l'action optimale a été évaluée à tord comme étant faible, en trop peu d'essais.

Une solution pour régler le problème serait de continuer de choisir les machines à tour de rôle pendant un certain temps, tant que les estimations associées ne sont pas stables. Une fois suffisamment de gains observés pour chaque machine, la prise de décisions présentée au point précédent pourrait être mise en place.

Question 3 (34 points sur 100)

Supposons les données suivantes en deux dimensions :

$$\mathbf{x}^1 = \begin{bmatrix} 2{,}50 \\ 1{,}00 \end{bmatrix}, \quad r^1 = 0, \quad \mathbf{x}^2 = \begin{bmatrix} 3{,}50 \\ 1{,}30 \end{bmatrix}, \quad r^2 = 0, \quad \mathbf{x}^3 = \begin{bmatrix} 2{,}00 \\ 2{,}00 \end{bmatrix}, \quad r^3 = 0,$$

$$\mathbf{x}^4 = \begin{bmatrix} 4{,}15 \\ 2{,}90 \end{bmatrix}, \quad r^4 = 1, \quad \mathbf{x}^5 = \begin{bmatrix} -0{,}10 \\ -1{,}50 \end{bmatrix}, \quad r^5 = 1, \quad \mathbf{x}^6 = \begin{bmatrix} -2{,}00 \\ -0{,}40 \end{bmatrix}, \quad r^6 = 1.$$

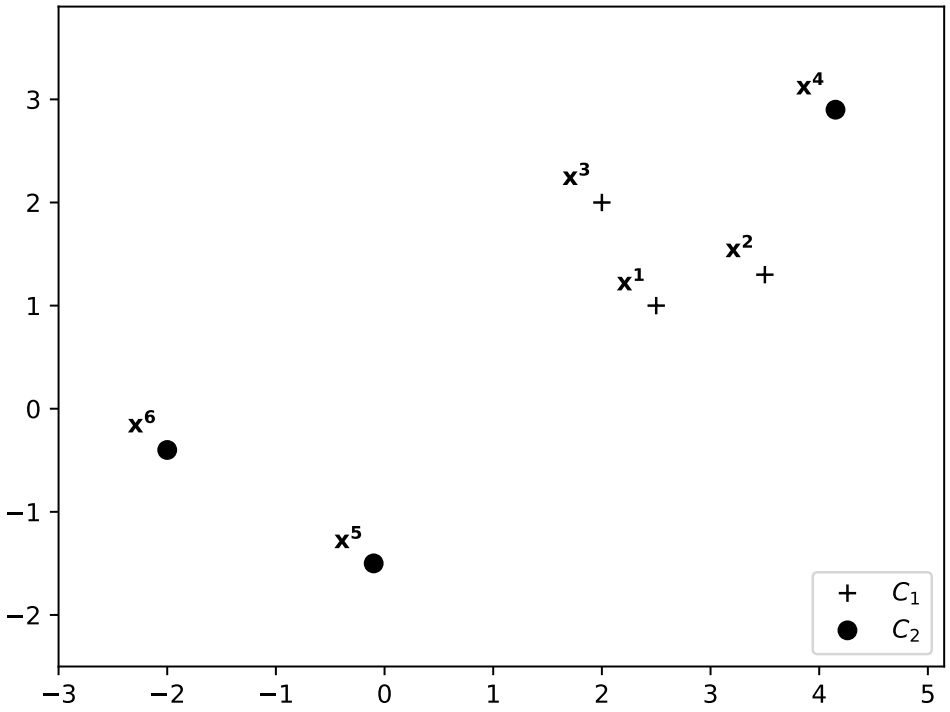
Le vecteur moyen $\mathbf{m}$ et la matrice de covariance $\mathbf{S}$ estimés de ces données sont les suivants :

$$\mathbf{m} = \begin{bmatrix} 1{,}67500 \\ 0.88333 \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} 5{,}37975 & 3{,}03150 \\ 3{,}03150 & 2{,}56567 \end{bmatrix}.$$

Les vecteurs propres et valeurs propres associés de cette matrice de covariance sont :

$$\lambda_1 = 7{,}3148, \quad \mathbf{c}_1 = \begin{bmatrix} 0{,}84291 \\ 0{,}53805 \end{bmatrix}, \quad \lambda_2 = 0{,}63059, \quad \mathbf{c}_2 = \begin{bmatrix} -0{,}53805 \\ 0{,}84291 \end{bmatrix}.$$

Finalement, les données sont tracées dans la figure suivante.



(10) (a) La procédure suivante a été brièvement présentée en classe afin d'initialiser l'algorithme K-means à partir d'une analyse en composantes principales (ACP) :

1. effectuer une ACP sur l'ensemble des données pour en extraire la composante principale (première composante) ;
2. projeter les données en une dimension selon la composante principale ;
3. partitionner les données en K groupes de taille égale dans l'espace unidimensionnel ;
4. en utilisant le partitionnement de l'étape précédente, calculer les centres des groupes dans l'espace d'origine.

Effectuez cette procédure pour déterminer les paramètres initiaux de l'algorithme K-means, avec $K = 2$ groupes.

Solution: La projection en une dimension se fait avec l'équation suivante :

$$z_1^t = \mathbf{c}_1^\top \mathbf{x}^t, t = 1, \ldots, 6,$$

où $\mathbf{c}_1$ est la première composante, soit celle avec la valeur propre associée la plus élevée ($\lambda_1 > \lambda_2$). La valeur des données projetées sur $\mathbf{c}_1$ est :

$$\begin{array}{lll} z_1^1 = 2{,}64533, & z_1^2 = 3{,}64966, & z_1^3 = 2{,}76193, \\ z_1^4 = 5{,}05843, & z_1^5 = -0{,}89137, & z_1^6 = -1{,}90105. \end{array}$$

Le tri de ces données en ordre croissant selon cette projection donne : z_1^6, z_1^5, z_1^1, z_1^3, z_1^2 et z_1^4. Le partitionnement en deux groupes de taille égale selon cet ordre est : $\{z_1^1, z_1^5, z_1^6\}$ et $\{z_1^2, z_1^3, z_1^4\}$, pour des assignations initiales correspondant à :

$$b^1(0) = 1, \quad b^2(0) = 2, \quad b^3(0) = 2, \quad b^4(0) = 2, \quad b^5(0) = 1, \quad b^6(0) = 1.$$

Les centres correspondants sont donc :

$$\mathbf{m}_1(0) = \begin{bmatrix} 0{,}13333 \\ -0{,}30000 \end{bmatrix}, \quad \mathbf{m}_2(0) = \begin{bmatrix} 3{,}21667 \\ 2.06667 \end{bmatrix}.$$

(8) (b) Exécutez l'algorithme K-means (avec $K = 2$ groupes) à partir l'initialisation obtenue à la sous-question précédente, jusqu'à convergence de l'algorithme.

Solution: En partant des centres $\mathbf{m}_1(0)$ et $\mathbf{m}_2(0)$ obtenues à la question précédente, nous recalculons les distances des données aux centres.

	$\|\mathbf{x}^t - \mathbf{m}_1(0)\|$	$\|\mathbf{x}^t - \mathbf{m}_2(0)\|$	$b^t(1)$
$\mathbf{x}^1$	2,70021	1,28506	2
$\mathbf{x}^2$	3,72752	0,81735	2
$\mathbf{x}^3$	2,96217	1,21849	2
$\mathbf{x}^4$	5,13552	1,25122	2
$\mathbf{x}^5$	1,22247	4,87046	1
$\mathbf{x}^6$	2,13568	5,77045	1

Les nouveaux centres correspondants sont :

$$\mathbf{m}_1(1) = \begin{bmatrix} -1{,}05 \\ -0{,}95 \end{bmatrix}, \quad \mathbf{m}_2(1) = \begin{bmatrix} 3{,}0375 \\ 1{,}8 \end{bmatrix}.$$

Comme les centres ont changé, on recalcule les distances et l'assignation aux centres.

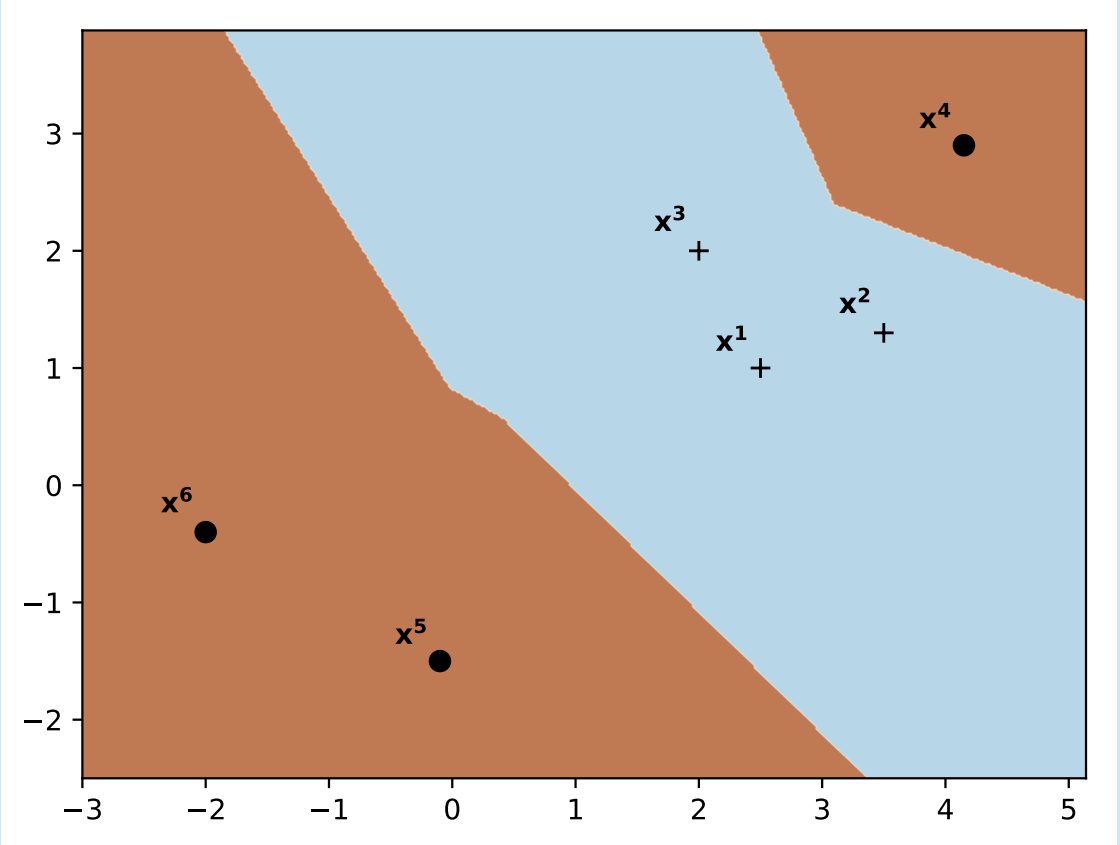
	$\|\mathbf{x}^t - \mathbf{m}_1(1)\|$	$\|\mathbf{x}^t - \mathbf{m}_2(1)\|$	$b^t(2)$
$\mathbf{x}^1$	4,05031	0,96380	2
$\mathbf{x}^2$	5,07592	0,68111	2
$\mathbf{x}^3$	4,24323	1,05660	2
$\mathbf{x}^4$	6,47012	1,56450	2
$\mathbf{x}^5$	1,09772	4,55345	1
$\mathbf{x}^6$	1,09772	5,49695	1

L'assignation aux centres $b^t(2)$ n'a pas changé donc l'algorithme a convergé. Les centres

$\mathbf{m}_1(1)$ et $\mathbf{m}_2(1)$ ainsi que l'assignation aux centres $b^t(1)$ correspond donc au résultat final.

(8) (c) Tracez les régions de décision selon les données d'entraînement pour un classifieur de type plus proche voisin (avec un seul voisin, $k = 1$). Tracez le tout dans votre **cahier bleu d'examen** (et non dans l'énoncé courant). Donnez également le taux de classement selon une méthodologie *leave-one-out* avec cette configuration, sur ces données.

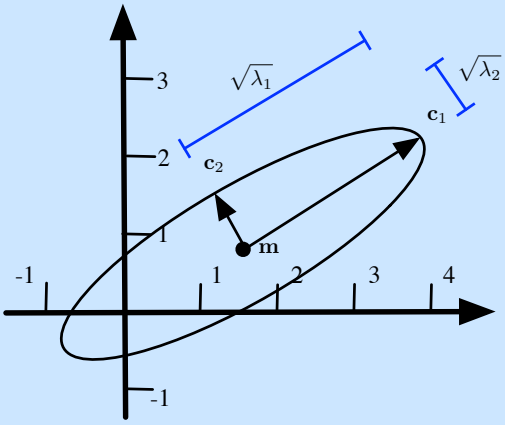
Solution: Les régions de décision avec un classifieur de type plus proche voisin sont les suivantes.



Le taux de classement selon une méthodologie *leave-one-out* sur ces données est de 83 %, comme la donnée $\mathbf{x}^4$ sera mal classée par son plus proche voisin, les autres données étant correctement classées.

(8) (d) Faites un graphique représentant la distribution des données en deux dimensions, en y traçant la courbe de contour correspondant à une distance de Mahanalobis de 1 (équivalent à une distance d'un écart-type en une dimension). Indiquez clairement dans le graphique l'utilisation des différentes valeurs données dans l'énoncé de la question, soit le vecteur moyen estimé $\mathbf{m}$, les valeurs propres λ_i et les directions des vecteurs propres $\mathbf{c}_i$.

Solution:



Question 4 (36 points sur 100)

Répondez aussi brièvement et clairement que possible aux questions suivantes.

(3) (a) Dans la formalisation mathématique présentée en classe, un modèle d'apprentissage supervisé est représenté par une fonction $\mathrm{h}(\mathbf{x}|\theta)$. Dans ce modèle, expliquez ce que représente les variables $\mathbf{x}$ et θ.

Solution: La variable $\mathbf{x}$ représente une donnée devant être traitée par le modèle.
La variable θ représente les paramètres du modèle qui sont appris sur un jeu d'entraînement.

(3) (b) Supposons que l'on fait l'évaluation de modèles de classement selon une approche de type validation croisée à K-plis (en anglais : K-*fold cross-validation*) avec un jeu de données $\mathcal{X}$. Donnez le nombre d'instances sur lequel le taux de classement a été calculé.

Solution: La taux de classement a été calculé sur l'ensemble des instances du jeux de données $\mathcal{X}$, soit N données selon la notation utilisée dans le cours.

(3) (c) Dans une approche de classement probabiliste, expliquez à quoi correspond $P(C_i|\mathbf{x})$

Solution: $P(C_i|\mathbf{x})$ correspond à la probabilité que la donnée $\mathbf{x}$ soit de la classe C_i selon le classement effectué.

(3) (d) D'après vous, où se situe un classifieur de type plus proche voisin (avec $k = 1$ voisin) selon le compromis biais-variance. Justifiez votre réponse.

Solution: Un classifieur de type plus proche voisin est un modèle de classement à biais nul mais à variance élevée. En effet, d'un jeu de données à l'autre, les frontières de décision peuvent varier significativement, impliquant ainsi une forte variance. Cependant, le modèle ne possède un biais faible dans la mesure qu'il peut représenter un phénomène

naturel sans imposer une forme de modèle, avec un résultat sur un très grand jeu de données convergeant vers le résultat bayésien optimal.

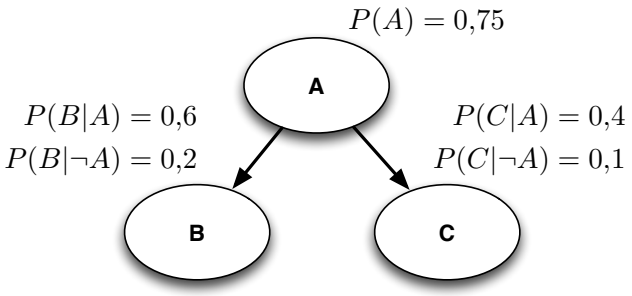
(3) (e) On dit que deux variables indépendantes ont une corrélation nulle, mais que l'inverse n'est pas nécessairement vrai. Expliquez cette affirmation.

Solution: Deux variables ayant une corrélation nulle peuvent avoir une dépendance non linéaire qui n'est pas capturée dans la mesure de corrélation.

(3) (f) Dans un contexte de classement paramétrique avec lois normales multivariées, donnez la condition sur les modèles faisant en sorte que les frontières de décision entre les classes soient linéaires.

Solution: Pour que les frontières soient linéaires, il faut que les valeurs des matrices de covariances soient les mêmes entre toutes les classes. Dit autrement, les matrices de covariance doivent être partagées entre les classes.

(3) (g) Soit le réseau bayésien suivant, calculez $P(C|\neg B)$.



Solution:

$$
\begin{aligned}
P(\neg B) &= P(\neg B|A)\,P(A) + P(\neg B|\neg A)\,P(\neg A) \\
&= (1-0.6) \times 0{,}75 + (1-0{,}2) \times (1-0{,}75) = 0{,}3 + 0{,}2 = 0{,}5 \\
P(A|\neg B) &= \frac{P(\neg B|A)\,P(A)}{P(\neg B)} = \frac{(1-0{,}6) \times 0{,}75}{0{,}5} = 0{,}6 \\
P(\neg A|\neg B) &= \frac{P(\neg B|\neg A)\,P(\neg A)}{P(\neg B)} = \frac{(1-0{,}2) \times (1-0{,}75)}{0{,}5} = 0{,}4 \\
P(C|\neg B) &= P(C|A,\neg B)\,P(A|\neg B) + P(C|\neg A,\neg B)\,P(\neg A|\neg B) \\
&= P(C|A)\,P(A|\neg B) + P(C|\neg A)\,P(\neg A|\neg B) \\
&= 0{,}4 \times 0{,}6 + 0{,}1 \times 0{,}4 = 0{,}24 + 0{,}04 = 0{,}28
\end{aligned}
$$

(3) (h) Expliquez précisément pourquoi les algorithmes de sélection de variables avant séquentielle et arrière séquentielle, tels que présentés dans le cours, sont des heuristiques sans garantie d'optimalité.

Solution: Ces algorithmes se basent sur des décisions basées sur une vue « myope » de la performance des variables potentiellement ajoutée, en utilisant l'interaction d'une variable à la fois avec un ensemble fixe de variables pour en évaluer la pertinence.

(3) (i) Expliquez l'effet d'une transformation blanchissante sur un jeu de données quelconque.

Solution: Une transformation blanchissante centre les données à l'origine et retire la covariance. L'application de la transformation blanchissante sur un jeu où les données suivent une loi normale multivariée quelconque ($\mathbf{x} \sim \mathcal{N}_D(\boldsymbol{\mu},\boldsymbol{\Sigma})$) transforme les données pour quelles suivent une loi normale multivariée de moyenne nulle et variance unitaire ($\mathbf{z} \sim \mathcal{N}_D(0,\boldsymbol{I})$).

(3) (j) Selon les explications faites en classe sur l'algorithme Espérance-Maximisation (EM), indiquez à quoi correspondent précisément les variables $\mathbf{z}^t$ ainsi que leur contenu.

Solution: Les variables $\mathbf{z}^t$ sont des variables cachées, non observables, donnant la véritable association des données $\mathbf{x}^t$ aux différents groupes $\{\mathcal{G}_i\}_{i=1}^K$. Ces variables sont des vecteurs binaires de taille K, avec une seule valeur à 1 et les autres valeurs à 0 (*one-hot vector*).

(3) (k) Expliquez pourquoi dit-on qu'il est préférable de bien normaliser les données avant l'application d'un classement de type k-plus proches voisins basé sur une distance euclidienne.

Solution: Avec la distance euclidienne, comme la plupart des distances, si le domaine des données pour une variable est beaucoup plus vaste qu'une les autres variables ($x_i \gg x_j, \forall j \neq i$), alors la distance sera essentiellement basée sur cette distance ($d(\mathbf{x},\mathbf{y}) \approx |x_i - y_i|$), ignorant ainsi de l'information pouvant se trouver dans les autres variables.

(3) (l) Expliquez pourquoi il est déconseillé de faire une condensation de Hart suivie d'une édition de Wilson, alors que le contraire est possible et même suggéré dans certains cas.

Solution: L'édition de Wilson vise essentiellement à retirer des données aberrantes ou bruitées du jeu de données, c'est-à-dire celles incohérentes avec les autres données selon un classement *leave-one-out*. La condensation de Hart vise à compresser les données en un ensemble de prototypes de taille minimale, contenant que les données essentielles pour le classement. En appliquant l'édition de Wilson après une condensation de Hart, on se trouverait à retirer beaucoup de prototypes comme ils seraient incohérents avec les autres actuellement en évaluation. Cependant, ceux-ci sont essentiels au jeu de prototypes selon une condensation de Hart. La performance qui en résulte risque d'être sérieusement affectée, comme des prototypes importants auront été éliminés.