Apprentissage et reconnaissance (GIF-4101 / GIF-7005)
Département de génie électrique et de génie informatique
Automne 2010

EXAMEN FINAL

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.

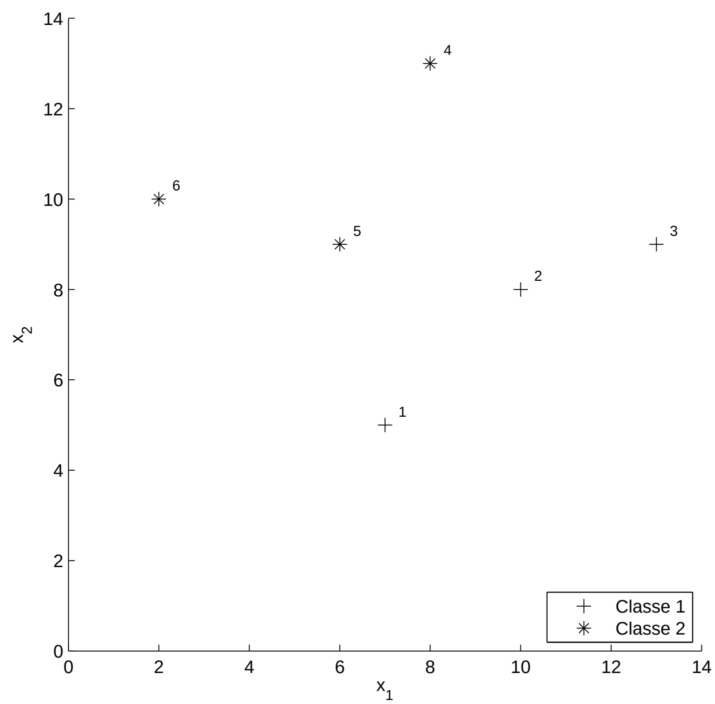
Question 1 (15 points sur 100)

Soit le jeu de données $\mathcal{X} = \{\mathbf{x}^t, r^t\}_{t=1}^6$ présenté ci-bas.

$$\mathbf{x}^1 = \begin{bmatrix} 7 \\ 5 \end{bmatrix}, \quad r^1 = -1, \quad \mathbf{x}^2 = \begin{bmatrix} 10 \\ 8 \end{bmatrix}, \quad r^2 = -1, \quad \mathbf{x}^3 = \begin{bmatrix} 13 \\ 9 \end{bmatrix}, \quad r^3 = -1,$$

$$\mathbf{x}^4 = \begin{bmatrix} 8 \\ 13 \end{bmatrix}, \quad r^4 = 1, \quad \mathbf{x}^5 = \begin{bmatrix} 6 \\ 9 \end{bmatrix}, \quad r^5 = 1, \quad \mathbf{x}^6 = \begin{bmatrix} 2 \\ 10 \end{bmatrix}, \quad r^6 = 1$$

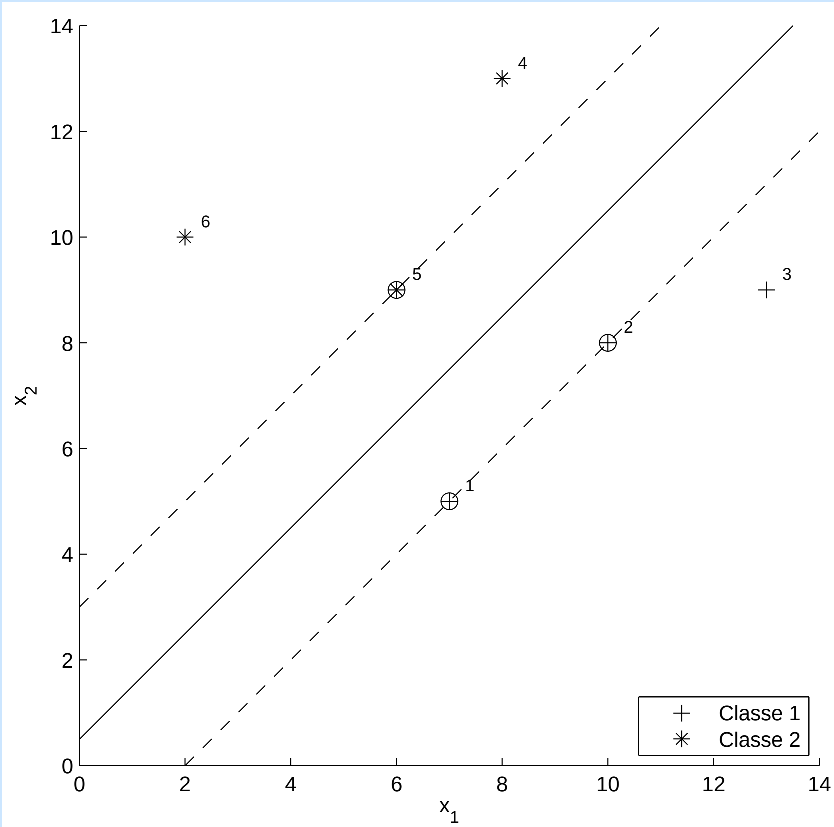
La figure suivante trace ces points en deux dimensions.



Supposons que l'on veut classer ces données avec un classifieur de type séparateur à vastes marges (SVM) utilisant un noyau linéaire ($K(\mathbf{x},\mathbf{x}') = \langle \mathbf{x},\mathbf{x}' \rangle$), sans marge floue.

(5) (a) Dans votre **cahier de réponse**, tracez les données du jeu $\mathcal{X}$, les marges géométriques maximales obtenues avec le SVM, l'hyperplan séparateur correspondant, et encerclez les données agissant comme vecteurs de support.

Solution:



(10) (b) Donnez les valeurs des poids $\mathbf{w}$ et biais w_0 correspondant au discriminant linéaire maximisant les marges géométriques tracées en a).
Indice : il n'est pas nécessaire de calculer les α^i pour répondre à la question.

Solution: Trois données sont identifiées comme vecteurs de support : $\mathbf{x}^1$, $\mathbf{x}^2$ et $\mathbf{x}^5$. En travaillant directement dans l'espace d'entrée, ceci nous donne trois équations et trois inconnus.

$$\begin{aligned}
\mathrm{h}(\mathbf{x}^1) &= w_2 x_2^1 + w_1 x_1^1 + w_0 = 5w_2 + 7w_1 + w_0 = -1 \\
\mathrm{h}(\mathbf{x}^2) &= w_2 x_2^2 + w_1 x_1^2 + w_0 = 8w_2 + 10w_1 + w_0 = -1 \\
\mathrm{h}(\mathbf{x}^5) &= w_2 x_2^5 + w_1 x_1^5 + w_0 = 9w_2 + 6w_1 + w_0 = 1
\end{aligned}$$

La résolution de ce système d'équation peut se faire par la suivante.

$$\begin{array}{rcl}
(8 \times \text{EQ1}) - (5 \times \text{EQ2}) & : & (40-40)w_2 + (56-50)w_1 + (8-5)w_0 = -8+5 \\
 & & 6w_1 + 3w_0 = -3 \\
(9 \times \text{EQ2}) - (8 \times \text{EQ3}) & : & (72-72)w_2 + (90-48)w_1 + (9-8)w_0 = -9-8 \\
 & & 42w_1 + w_0 = -17 \\
\text{L1} - (3 \times \text{L2}) & : & (6-126)w_1 + (3-3)w_0 = -3+51 \\
 & & -120w_1 = 48 \qquad \Rightarrow \qquad w_1 = -0{,}4 \\
\text{L1} - (6 \times \text{L3}) & : & (6-6)w_1 + (3-0)w_0 = -3-6(-0{,}4) \\
 & & 3w_0 = -3+2{,}4 \qquad \Rightarrow \qquad w_0 = -0{,}2 \\
\text{EQ1} - (7 \times \text{L3}) - \text{L4} & : & 5w_2 + (7-7)w_1 + (1-1)w_0 = -1-7(-0{,}4)-(-0{,}2) \\
 & & 5w_2 = -1+2{,}8+0{,}2 = 2 \qquad \Rightarrow \qquad w_2 = 0{,}4
\end{array}$$

La résolution de ce système d'équations nous donne les valeurs suivantes :

$$w_2 = 0{,}4,\ w_1 = -0{,}4,\ w_0 = -0{,}2.$$

Question 2 (10 points sur 100)

Soit le classifieur à noyau suivant, utilisant le critère d'erreur du perceptron.

$$\begin{array}{rcl}
\mathrm{h}(\mathbf{x}|\boldsymbol{\alpha},w_0) & = & \displaystyle\sum_{\mathbf{x}^s \in \mathcal{X}} \alpha^s r^s K(\mathbf{x}^s,\mathbf{x}) + w_0, \\
\alpha^t & \geq & 0 \quad \forall t, \\
\text{avec :} & & \\
E(\boldsymbol{\alpha},w_0|\mathcal{X}) & = & \displaystyle -\sum_{\mathbf{x}^t \in \mathcal{Y}} r^t \mathrm{h}(\mathbf{x}^t|\boldsymbol{\alpha},w_0) + \lambda \frac{1}{2} \sum_{\alpha^t \in \boldsymbol{\alpha}} (\alpha^t)^2, \\
\mathcal{Y} & = & \{\mathbf{x}^t \in \mathcal{X} | r^t \mathrm{h}(\mathbf{x}^t|\boldsymbol{\alpha},w_0) < 0\}.
\end{array}$$

Donnez les équations pour mettre à jour les $\boldsymbol{\alpha}$ et w_0 selon une descente du gradient.

Solution:

$$\begin{aligned}
\frac{\partial E}{\partial \alpha^i} &= \frac{\partial}{\partial \alpha^i}\left[-\sum_{\mathbf{x}^t\in\mathcal{Y}} r^t \left(\sum_{\mathbf{x}^s\in\mathcal{X}} \alpha^s r^s K(\mathbf{x}^s,\mathbf{x}^t) + w_0\right) + \lambda\frac{1}{2}\sum_{\alpha^t\in\boldsymbol{\alpha}} (\alpha^t)^2\right] \\
&= -\sum_{\mathbf{x}^t\in\mathcal{Y}} r^t r^i K(\mathbf{x}^i,\mathbf{x}^t) + \lambda\alpha^i \\
\Delta\alpha^i &= -\eta\frac{\partial E}{\partial \alpha^i} = \eta\left[\sum_{\mathbf{x}^t\in\mathcal{Y}} r^t r^i K(\mathbf{x}^i,\mathbf{x}^t) - \lambda\alpha^i\right] \\
\frac{\partial E}{\partial w_0} &= \frac{\partial}{\partial w_0}\left[-\sum_{\mathbf{x}^t\in\mathcal{Y}} r^t \left(\sum_{\mathbf{x}^s\in\mathcal{X}} \alpha^s r^s K(\mathbf{x}^s,\mathbf{x}^t) + w_0\right) + \lambda\frac{1}{2}\sum_{\alpha^t\in\boldsymbol{\alpha}} (\alpha^t)^2\right] \\
&= -\sum_{\mathbf{x}^t\in\mathcal{Y}} r^t \\
\Delta w_0 &= -\eta\frac{\partial E}{\partial w_0} = \eta\sum_{\mathbf{x}^t\in\mathcal{Y}} r^t
\end{aligned}$$

Question 3 (15 points sur 100)

Dans le cours, il a été avancé qu'un perceptron multi-couches (PMC) avec plusieurs couches de neurones et utilisant une fonction de transfert linéaire ($\mathrm{f}_{lin}(a) = a$) pour tous les neurones peut être simplifié comme PMC à une seule couche de neurones avec fonction de transfert linéaire. Démontrez que cette affirmation est vraie à partir des équations du PMC modélisant la propagation des données de l'entrée vers la sortie.

Solution: Pour répondre à cette question, on va d'abord démontrer qu'un PMC à deux couches avec fonction de transfert linéaire peut être simplifié en un PMC à une couche. L'équation décrivant la propagation des données de l'entrée vers la sortie y_k^t dans un neurone k sur la couche l de la donnée $\mathbf{x}^t$ est la suivante :

$$y_k^t = \mathrm{f}\left(\sum_{j=1}^{R} w_{k,j} y_j^t + w_{k,0}\right),$$

où y_j^t représente les sorties des neurones de la couche précédente, soit la couche $l-1$, pour la donnée. Avec une fonction de transfert linéaire la fonction se simplifie par ce qui suit :

$$y_k^t = \sum_{j=1}^{R} w_{k,j} y_j^t + w_{k,0}.$$

Si on substitue la même équation pour les sortie y_j^t de la couche k, on obtient l'équation suivante :

$$\begin{aligned} y_k^t &= \sum_{j=1}^{R} w_{k,j} \left(\sum_{i=1}^{Q} w_{j,i} y_i^t + w_{j,0} \right) + w_{k,0} \\ &= \left(\sum_{j=1}^{R} w_{k,j} \sum_{i=1}^{Q} w_{j,i} y_i^t \right) + \left(\sum_{j=1}^{R} w_{k,j} w_{j,0} \right) + w_{k,0} \\ &= \sum_{i=1}^{Q} \left(\sum_{j=1}^{R} w_{k,j} w_{j,i} \right) y_i^t + \left(\sum_{j=1}^{R} w_{k,j} w_{j,0} + w_{k,0} \right), \end{aligned}$$

où y_i^t représente la sortie de la couche $l-2$, soit la couche précédent la couche du neurone j. On peut alors poser les variables suivantes :

$$\begin{aligned} w'_{k,i} &= \sum_{j=1}^{R} w_{k,j} w_{j,i}, \\ w'_{k,0} &= \sum_{j=1}^{R} w_{k,j} w_{j,0} + w_{k,0}, \end{aligned}$$

ce qui nous permet de réexprimer l'équation pour le neurone k directement en fonction des neurone des couche $l-2$:

$$y_k^t = \sum_{i=1}^{Q} w'_{k,i} y_i^t + w'_{k,0}.$$

De cette façon on vient de simplifier les couches l et $l-1$ en une seule couche de neurones linéaires connectés directement sur la sortie des neurones de la couche $l-2$.

Par induction, on peut ainsi simplifer successivement toutes les couches du PMC deux à deux, pour obtenir un réseau simplifié avec une seule couche de neurones avec fonction de tranfert linéaire.

Question 4 (15 points sur 100)

La fonction $\mathrm{h}_{j,i}$ correspond à la décision du classifieur h_j concernant la classe C_i. Ainsi, pour un problème à trois classes, les fonctions $\mathrm{h}_{j,1}$, $\mathrm{h}_{j,2}$ et $\mathrm{h}_{j,3}$ retournent les valeurs de la décision du classifieur h_j pour les classes C_1, C_2 et C_3, respectivement.

(5) (a) Supposons que l'on bâtit un ensemble de $L = 5$ classifieurs prenant des décisions sur $K = 3$ classes. La fonction $\bar{\mathrm{h}}_i$ retourne le résultat de l'ensemble pour la classe C_i. Pour

une donnée $\mathbf{x}$ particulière, on obtient les résultats suivants avec cet ensemble.

	C_1	C_2	C_3
$\mathrm{h}_{1,i}(\mathbf{x})$	0,3	0,5	0,45
$\mathrm{h}_{2,i}(\mathbf{x})$	0,4	0,4	0,35
$\mathrm{h}_{3,i}(\mathbf{x})$	0,45	0,6	0,5
$\mathrm{h}_{4,i}(\mathbf{x})$	0,1	0,2	0,15
$\mathrm{h}_{5,i}(\mathbf{x})$	0,3	0,2	0,3

Pour chacune des fonctions de combinaison suivantes, utilisées pour calculer le résultat de la fonction de l'ensemble par classe $\bar{\mathrm{h}}_i$, donnez les décisions de classement pour la donnée $\mathbf{x}$:

1. somme ;
2. maximum ;
3. médiane.

Solution: Les résultats pour chaque fonction de combinaison sont les suivants :

	C_1	C_2	C_3	décision
$\mathrm{h}_{1,i}(\mathbf{x})$	0,3	0,5	0,45	
$\mathrm{h}_{2,i}(\mathbf{x})$	0,4	0,4	0,35	
$\mathrm{h}_{3,i}(\mathbf{x})$	0,45	0,6	0,5	
$\mathrm{h}_{4,i}(\mathbf{x})$	0,1	0,2	0,15	
$\mathrm{h}_{5,i}(\mathbf{x})$	0,3	0,2	0,3	
$\bar{\mathrm{h}}_i(\mathbf{x}) = \sum_{j=1}^{L} \mathrm{h}_{j,i}(\mathbf{x})$	1,55	1,9	1,75	C_2
$\bar{\mathrm{h}}_i(\mathbf{x}) = \max_{j=1}^{L} \mathrm{h}_{j,i}(\mathbf{x})$	0,45	0,6	0,5	C_2
$\bar{\mathrm{h}}_i(\mathbf{x}) = \mathrm{med}_{j=1}^{L} \mathrm{h}_{j,i}(\mathbf{x})$	0,3	0,4	0,35	C_2

.

(5) (b) Supposons maintenant que les décisions sont binaires, de sorte que les valeurs pour chaque classe des classifieurs de l'ensemble sont $\mathrm{h}_{j,i}(\mathbf{x}) \in \{0,1\}$, et que l'on obtienne les résultats suivants pour la donnée $\mathbf{x}$ pour un ensemble de $L = K = 3$ classifieurs.

	C_1	C_2	C_3
$\mathrm{h}_{1,i}(\mathbf{x})$	0	1	1
$\mathrm{h}_{2,i}(\mathbf{x})$	0	1	0
$\mathrm{h}_{3,i}(\mathbf{x})$	0	0	1

Si cet ensemble est entraîné selon une approche revenant à *un contre tous*, avec la matrice de décision suivante,

$$\mathbf{W} = \begin{bmatrix} +1 & -1 & -1 \\ -1 & +1 & -1 \\ -1 & -1 & +1 \end{bmatrix},$$

donnez la décision de classement pour cette donnée.

Solution: La décision de l'ensemble pour la donnée est calculé selon l'équation

$$\bar{\mathrm{h}}_i(\mathbf{x}) = \sum_j w_{i,j} \mathrm{h}_{j,i}(\mathbf{x}).$$

Pour la donnée $\mathbf{x}$, le résultat est $\bar{\mathrm{h}}_1(\mathbf{x}) = 0$, $\bar{\mathrm{h}}_2(\mathbf{x}) = 0$ et $\bar{\mathrm{h}}_3(\mathbf{x}) = 0$. Il y a donc ambiguïté, la donnée ne peut donc pas être assignée à une classe.

(5) (c) Supposons maintenant que l'on veut utiliser une approche avec code de correction d'erreur comportant huit classifieurs. On a déjà établi les sept première colonnes de la matrice de décision $\mathbf{W}$ comme suit :

$$\mathbf{W} = \begin{bmatrix} +1 & -1 & -1 & +1 & +1 & -1 & +1 & ? \\ -1 & +1 & -1 & +1 & -1 & +1 & -1 & ? \\ -1 & -1 & +1 & -1 & +1 & +1 & +1 & ? \end{bmatrix}.$$

Déterminez les valeurs sur la dernière colonnes de la matrice de décision permettant de tolérer jusqu'à deux erreurs par les classifieurs de base.

Solution: Selon cette matrice, les distances de Hamming entre les différentes lignes sont les suivantes :
- Ligne 1 vs 2 : distance de 5 ;
- Ligne 1 vs 3 : distance de 4 ;
- Ligne 2 vs 3 : distance de 5.

Il faut donc augmenter la distance de Hamming entre les lignes 1 et 3. À cette fin, on doit donc faire en sorte que les lignes 1 et 3 aient une valeur différente pour le neuvième colonne. Ceci nous permettre d'avoir une distance de Hamming d'au moins 5 entre chaque ligne, ce qui nous permet de tolérer $\lfloor (5-1)/2 \rfloor = 2$ erreurs. Une solution possible pour compléter la matrice de décision est donc :

$$\mathbf{W} = \begin{bmatrix} +1 & -1 & -1 & +1 & +1 & -1 & +1 & +1 \\ -1 & +1 & -1 & +1 & -1 & +1 & -1 & -1 \\ -1 & -1 & +1 & -1 & +1 & +1 & +1 & -1 \end{bmatrix}.$$

Une alternative valable (parmi plusieurs autres) serait :

$$\mathbf{W} = \begin{bmatrix} +1 & -1 & -1 & +1 & +1 & -1 & +1 & -1 \\ -1 & +1 & -1 & +1 & -1 & +1 & -1 & -1 \\ -1 & -1 & +1 & -1 & +1 & +1 & +1 & +1 \end{bmatrix}.$$

Question 5 (45 points sur 100)

Répondez aussi brièvement et clairement que possible aux questions suivantes.

(3) (a) Dans les séparateurs à vaste marge (SVM), utilisant la formulation avec marge douce, que

représente une valeur de la variable *slack* ξ^t lorsqu'elle est comprise dans $0 < \xi^t \leq 1$?

Solution: Une valeur $\xi^t \in]0{,}1]$ indique que la donnée $\mathbf{x}^t$ associée à la variable est bien classée par le SVM (du bon côté de l'hyperplan séparateur), mais qu'elle est dans la marge.

(3) (b) Avec l'analyse en composantes principales à noyau, on effectue une extraction des vecteurs propres et valeurs propres de la matrice de Gram normalisée. Indiquez combien de valeurs comporte la première composante principale extraite à l'aide de cette approche.

Solution: La première composante principale comporte N données, soit le nombre de données comprises dans le jeu de données et utilisées pour calculer la matrice de Gram.

(3) (c) Indiquez précisemment quel élément de l'algorithme de rétropropagation des erreurs du perceptron multi-couches fait en sorte qu'on le qualifie d'algorithme stochastique.

Solution: Le perceptron multi-couche est un algorithme stochastique car les poids initiaux du réseau de neurones sont initialisés aléatoirement.

(3) (d) Dans l'algorithme de rétropropagation des erreurs du perceptron multi-couches, on utilise la règle de chaînage des dérivées pour calculer la correction à appliquer aux poids et biais du réseau. Par exemple, la correction de poids sur une couche de sortie se calcule à partir des dérivées partielles données dans l'équation suivante,

$$\frac{\partial E^t}{\partial w_{j,i}} = \frac{\partial E^t}{\partial e_j^t}\frac{\partial e_j^t}{\partial y_j^t}\frac{\partial y_j^t}{\partial a_j^t}\frac{\partial a_j^t}{\partial w_{j,i}}.$$

Quelle est la valeur de la dérivée partielle $\partial a_j^t / \partial w_{j,i}$ dans cette équation.

Solution: La dérivée partielle $\frac{\partial a_j^t}{\partial w_{j,i}}$ représente la valeur de la sortie du sommateur du neurone, avant qu'il soit appliqué comme entrée à la fonction de transfert, selon le poids $w_{j,i}$, relié à la sortie y_i de la couche précédente. Cette dérivée partielle est calculée comme suit :

$$\frac{\partial a_j^t}{\partial w_{j,i}} = \frac{\partial}{\partial w_{j,i}} \sum_{i=1}^{R} w_{j,i} y_i^t + w_{j,0} = y_i^t.$$

Sa valeur est donc y_i^t, soit la sortie du neurone i de la couche précédente pour la donnée $\mathbf{x}^t$.

(3) (e) Dans les méthodes par ensemble, il est démontré que la variance des performance d'un ensemble $\bar{\mathrm{h}}$ formé de L classifieurs individuels h_j décroît selon la taille de l'ensemble selon

$$\mathrm{Var}(\bar{\mathrm{h}}) = \frac{1}{L}\mathrm{Var}(\mathrm{h}_j).$$

Cette formulation fait cependant que la réponse des classifieurs individuels h_j respecte l'hypothèse iid (identiquement et indépendamment distribués). Indiquez en vos propres mots ce que signifie cette hypothèse dans le contexte présent de classement avec ensembles.

Solution: L'hypothèse iid sur la réponse des classifieurs individuels signifie qu'il n'y a aucune correlation entre les sorties des classifieurs, de sorte qu'un classifieur h_i n'a pas plus de chance de se tromper sur une donnée $\mathbf{x}$ si un autre classifieur fait une erreur de classement sur cette donnée ou non. De cette façon, les réponses, et donc les erreurs de classement, sont indépendantes d'un classifieur à l'autre, ce qui est une hypothèse forte.

(3) (f) Avec l'algorithme AdaBoost, on modifie une probabilité p_j^t qu'une donnée $\mathbf{x}^t$ soit échantillonnée pour entraîner un classifieur à l'itération j. Indiquez de quelle façon, d'un point de vue conceptuel, cette probabilité est modifiée à chaque itération de l'algorithme.

Solution: La probabilité p_j^t est augmentée à une itération particulière si la donnée $\mathbf{x}^t$ est mal classée par le classifieur h_j généré à cette itération. Par la normalisation des probabilités, la valeur de p_j^t se trouve également à être diminuée si la donnée est bien classée par le classifieur.

(3) (g) Lorsque l'on fait des expérimentations avec des algorithmes d'apprentissage supervisé pour faire du classement, on effectue souvent du partitionnement des ensembles de données avec **stratification**, où l'on respecte les proportions des données selon les différentes classes du problème (probabilités *a priori*). Indiquez pourquoi cette approche est souhaitable dans un contexte d'expérimentation et d'analyse, comparativement à un partitionnement sans stratification.

Solution: Le partitionnement avec stratification permet d'éliminer certaines variations dans la performances des algorithmes de classement liées à la proportion des données selon les classes. En effet, certains algorithmes peuvent être sensibles à la balance entre les classes dans les jeux de données, qui est un facteur incontrôlable, dont on veut minimiser l'impact sur les résultats.

(3) (h) Dans le test de l'Analyse de variance (ANOVA), on veut comparer plusieurs algorithmes de classement, en tenant de vérifier l'hypothèse H_0 à l'effet que les moyennes des performances μ_j pour chaque classifieur sont égales. Pour vérifier cette hypothèse, on calcule deux estimateurs σ^2 de la variance des résultats pour chaque classifieur. Indiquez clairement ce que sont chacun de ces estimateurs de la variance et de quelle façon on les utilise pour déterminer si l'hypothèse H_0 est valide.

Solution: Le premier estimateur de la variance σ^2 est calculé en supposant que l'hypothèse H_0 est valide, soit que les moyennes des performances μ_j pour chaque classifieur sont égales. Le deuxième estimateur de la variance σ^2 est estimé en ignorant cette hypothèse, supposant que les moyennes des performances μ_j peuvent différer d'un classifieur à l'autre. Ainsi, si le premier estimateur σ^2 est statistiquement identique au deuxième estimateur σ^2, l'hypothèse H_0 s'avère être vérifiée, autrement elle est invalidée.

(3) (i) Indiquez précisemment les variables formant le modèle λ d'un modèle de Markov caché.

Solution: Un modèle de Markov caché est formé des variables $\lambda = \{\mathbf{A},\mathbf{B},\mathbf{\Pi}\}$, soit les probabilités de transitions entre états $\mathbf{A}$, les probabilités d'observations $\mathbf{B}$ et les probabilités d'état initial $\mathbf{\Pi}$.

(3) (j) Selon une méthode d'évaluation des performances de type *leave-one-out*, indiquez combien de fois une données particulière $\mathbf{x}^t \in \mathcal{X}$ de l'ensemble sera utilisée pour entraîner le classifieur évalué.

Solution: La donnée $\mathbf{x}^t$ sera utilisée dans pour l'entraînement $N-1$ fois, étant utilisé fois supplémentaire comme donnée de test, où N est le nombre de données dans l'ensemble $\mathcal{X}$.

(3) (k) Dans le problème d'évaluation avec un modèle de Markov caché, on veut évaluer la probabilité $P(O|\lambda)$ d'avoir un certain séquence d'observations O avec le modèle λ. Cette probabilité peut se calculer selon l'équation suivante,

$$P(O|\lambda) = \sum_{\forall S} P(O,S|\lambda).$$

Cependant, le calcul de cette probabilité n'est pas tractable, computationnellement parlant. Indiquez comment on doit procéder pour évaluer la probabilité $P(O|\lambda)$ selon la méthode vue en classe, qui comporte une complexité algorithmique raisonnable.

Solution: La procédure avant permet de calculer la probabilité $P(O|\lambda)$ en stockant des résultats intermédiaires dans une variable $\alpha_t(i) \equiv P(\{o_1,\ldots,o_t\},s_t = S_i|\lambda)$, représentant la probabilité d'observer la partie de la séquence allant jusqu'au temps t, en supposant que l'on est dans l'état S_i. On peut calculer les $\alpha_t(i)$ récursivement, selon les valeurs $\alpha_{t-1}(j)$, de sorte que les calculs seront considérablement réduits.

(3) (l) L'algorithme Baum-Welch permet de calculer le modèle λ d'un modèle de Markov caché à partir d'observations. Cet algorithme implique le calcul d'une probabilité $\xi_t(i,j)$. Indiquez ce que signifie précisemment cette probabilité.

Solution: La probabilité $\xi_t(i,j) \equiv P(s_t = S_i,s_{t+1} = S_j|O,\lambda)$ représente la probabilité d'être dans l'état S_i au temps t et d'être ensuite dans l'état S_j au temps $t+1$, étant donné le modèle λ et une séquence particulière d'observations O.

(3) (m) Indiquez ce que représente précisément $Q(s,a)$ dans un contexte d'apprentissage par renforcement.

Solution: $Q(s,a)$ représente la fonction de valeur d'action, soit l'estimation de l'espérance de la récompense à long terme associée à effectuer l'action a dans l'état s.

(3) (n) Selon Bellman, le valeur d'une action dans un certain état est donnée par :

$$Q^*(s,a) = \sum_{s'} \mathcal{P}^a_{ss'} \left[\mathcal{R}^a_{ss'} + \gamma \max_{a'} Q^*(s',a') \right] .$$

Indiquez pourquoi il n'est pas possible d'appliquer directement cette équation dans le contexte où l'on ne possède pas de modèle satisfaisant de l'environnement.

Solution: Cette équation requiert de connaître les probabilités de transitions $\mathcal{P}^a_{ss'}$ entre l'état s et s' lorsque l'on effectue l'action a. Pour connaître ces probabilités de transitions, on doit avoir un modèle précis de l'environnement. De même, on doit également connaître la récompense immédiate espérée $\mathcal{R}^a_{ss'}$, qui elle aussi requiert avoir un modèle de l'environnement.

(3) (o) Expliquez de quelle façon on détermine les actions effectuées par un agent dans un contexte d'apprentissage par renforcement, lorsque l'agent utilise une politique dite ϵ-greedy.

Solution: La politique ϵ-greedy consiste à choisir une action au hasard selon une probabilité ϵ, et autrement choisir l'action optimale selon les valeurs d'action $Q(s,a)$ actuelles.