Centre Universitaire de Mila	Examen de rattrapage en « Apprentissage Automatique »
Institut : ST, Département : MI	Durée de l'examen : 1h30 / Mila le : 21-09-2019
Master 1 – STIC	Documents non autorisés

Exercice 1 : SVM (08 points)

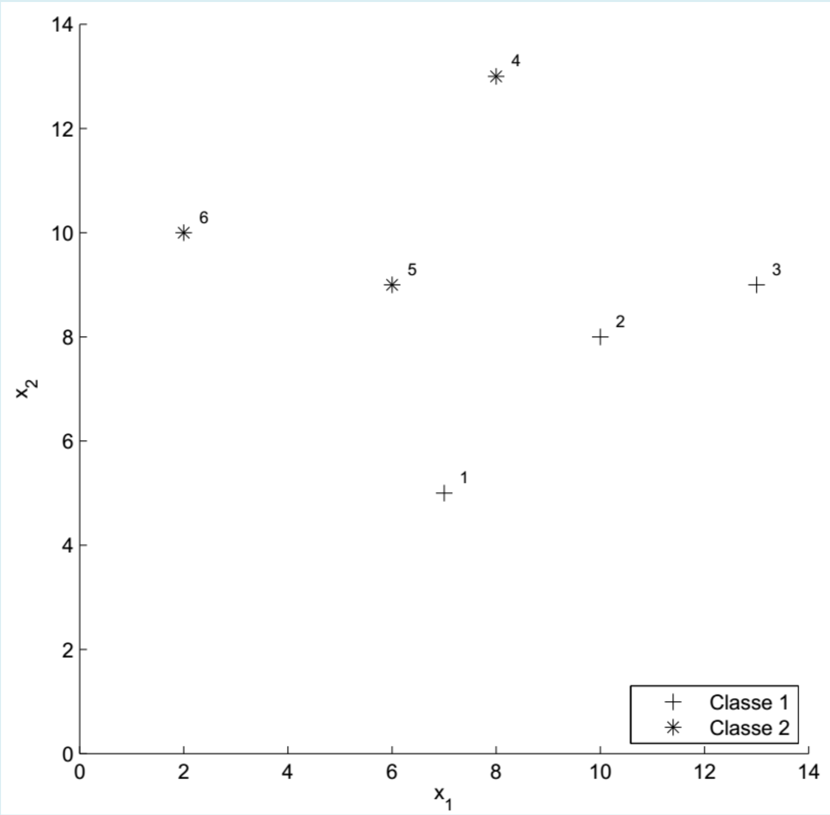
Soit l'ensemble de données $\mathbf{X} = \{(\mathbf{x}^{(\mathbf{t})}, \mathbf{y}^{(\mathbf{t})}),\ \boldsymbol{t} = \mathbf{1}, .., \mathbf{6}\}$ présenté ci-bas.

$$\mathbf{x}^{(1)} = \begin{bmatrix} 7 \\ 5 \end{bmatrix}, \mathbf{y}^{(1)} = -1, \qquad \mathbf{x}^{(2)} = \begin{bmatrix} 10 \\ 8 \end{bmatrix}, \mathbf{y}^{(2)} = -1, \qquad \mathbf{x}^{(3)} = \begin{bmatrix} 13 \\ 9 \end{bmatrix}, \mathbf{y}^{(3)} = -1,$$

$$\mathbf{x}^{(4)} = \begin{bmatrix} 8 \\ 13 \end{bmatrix}, \mathbf{y}^{(4)} = 1, \qquad \mathbf{x}^{(5)} = \begin{bmatrix} 6 \\ 9 \end{bmatrix}, \mathbf{y}^{(5)} = 1, \qquad \mathbf{x}^{(6)} = \begin{bmatrix} 2 \\ 10 \end{bmatrix}, \mathbf{y}^{(6)} = 1$$

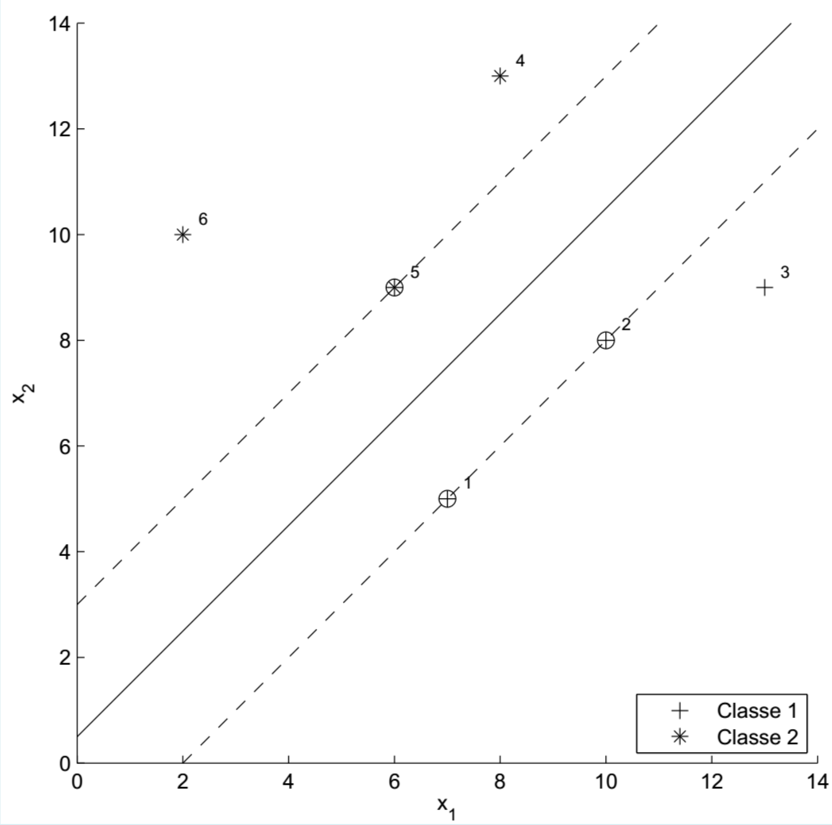
1) Tracer ces points en deux dimensions.

Solution



2) Supposons que l'on veut classer ces données avec un classifieur de type SVM utilisant un noyau linéaire ($\boldsymbol{K}(\mathbf{x}, \mathbf{x}') = \langle \mathbf{x}, \mathbf{x}' \rangle$) Tracez les données de l'ensemble $\mathbf{X}$, les marges géométriques maximales obtenues avec le SVM, l'hyperplan séparateur correspondant, et encerclez les données agissant comme vecteurs de support.

Solution



3) Donnez les valeurs des poids $\boldsymbol{w}$ et biais $\boldsymbol{w_0}$ correspondant au discriminant linéaire maximisant les marges géométriques tracées en question **2)**.

Solution

Trois données sont identifiées comme vecteurs de support : x_1, x_2 et x_5. En travaillant directement dans l'espace d'entrée, ceci nous donne trois équations et trois inconnus.

$$\mathbf{h(x^1) = w_2}x_2^1 + \mathbf{w_1}x_1^1 + \mathbf{w_0} = \mathbf{5w_2 + 7w_1 + w_0 = -1}$$

$$\mathbf{h(x^2) = w_2}x_2^2 + \mathbf{w_1}x_1^2 + \mathbf{w_0} = \mathbf{8w_2 + 10w_1 + w_0 = -1}$$

$$\mathbf{h\left(x^5\right) = w_2}x_2^5 + \mathbf{w_1}x_1^5 + \mathbf{w_0} = \mathbf{9w_2 + 6w_1 + w_0 = 1}$$

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

$$\mathbf{(8 \times EQ1) - (5 \times EQ2) : (40 - 40)w_2 + (56 - 50)w_1 + (8 - 5)w_0 = -8 + 5}$$

$$\mathbf{6w_1 + 3w_0 = -3}$$

$$\mathbf{(9 \times EQ2) - (8 \times EQ3) : (72 - 72)w_2 + (90 - 48)w_1 + (9 - 8)w_0 = -9 - 8}$$

$$\mathbf{42w_1 + w_0 = -17}$$

$$\mathbf{L1 - (3 \times L2) : (6 - 126)w_1 + (3 - 3)w_0 = -3 + 51}$$

$$\mathbf{-120w_1 = 48 \Rightarrow w_1 = -0.4}$$

$$\boldsymbol{L1 - (6 \times L3) : (6 - 6)w_1 + (3 - 0)w_0 = -3 - 6(-0.4)}$$

$$\mathbf{3w_0 = -3 + 2.4 \Rightarrow w_0 = -0.2}$$

$$\mathbf{EQ1 - (7 \times L3) - L4 : 5w_2 + (7 - 7)w_1 + (1 - 1)w_0}$$

$$\mathbf{= -1 - 7(-0.4) - (-0.2)}$$

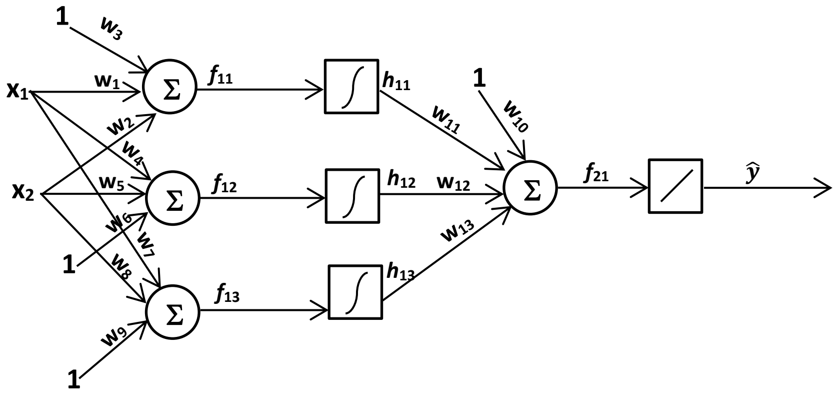
$$5w_2 = -1 + 2.8 + 0.2 = 2 \Rightarrow w_2 = 0.4$$

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

$$w_2 = 0.4, w_1 = -0.4, w_0 = -0.2$$

<u>Exercice 2 :</u> Réseaux de neurones (12 points)

Soit le réseau de neurones multicouches décrit par le graphe suivant :



1- Donner les formules mathématiques qui déterminent les sorties intermédiaires f_{11}, f_{12}, f_{13}, h_{11}, h_{12}, h_{13}, f_{21} ainsi que la sortie finale $\hat{y}$.

Solution : Propagation en avant (forward propagartion)

- $f_{11} = w_3 + w_1x_1 + w_2x_2$
- $f_{12} = w_6 + w_4x_1 + w_5x_2$
- $f_{13} = w_9 + 7x_1 + w_8x_2$
- $h_{11} = sigm(f_{11}) = \frac{1}{1+e^{-f_{11}}}$
- $h_{12} = sigm(f_{12}) = \frac{1}{1+e^{-f_{12}}}$
- $h_{13} = sigm(f_{13}) = \frac{1}{1+e^{-f_{13}}}$
- $\hat{y} = f_{21} = w_{10} + w_{11}h_{11} + w_{12}h_{12} + w_{13}h_{13}$

2- Soit la fonction d'erreur : $\mathbf{E}(\mathbf{w}) = (y - \hat{y})^2$

En appliquant l'algorithme de propagation en arrière (backpropagation), trouver les expressions des mises à jour des paramètres Δw_j pour **j = 1,…,13.**

Solution : Propagation en arrière (backpropagartion algorithm)

La fonction d'erreur est donnée par $\mathrm{E}(\mathrm{w}) = (y - \hat{y})^2$

Donc, on aura $\frac{\partial E(w)}{\partial w_j} = -2(y - \hat{y})\frac{\partial \hat{y}}{\partial w_j}$

D'après la propagation en avant, on a : $\hat{y} = f_{21} = w_{10} + w_{11}h_{11} + w_{12}h_{12} + w_{13}h_{13}$

Donc, les dérivées $\frac{\partial \hat{y}}{\partial w_j}$ **peuvent être calculées par :**

- $\frac{\partial \hat{y}}{\partial w_1} = \frac{\partial \hat{y}}{\partial h_{11}}\frac{\partial h_{11}}{\partial f_{11}}\frac{\partial f_{11}}{\partial w_1} = w_{11}h_{11}(1 - h_{11})x_1$
- $\frac{\partial \hat{y}}{\partial w_2} = \frac{\partial \hat{y}}{\partial h_{11}}\frac{\partial h_{11}}{\partial f_{11}}\frac{\partial f_{11}}{\partial w_2} = w_{11}h_{11}(1 - h_{11})x_2$

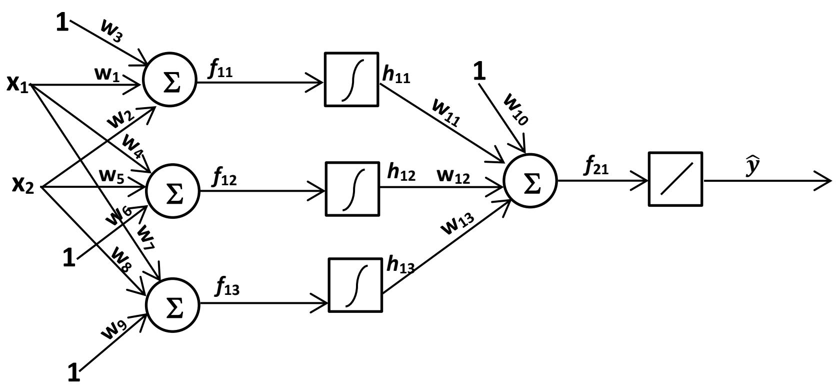
- $\frac{\partial \hat{y}}{\partial w_3} = \frac{\partial \hat{y}}{\partial h_{11}} \frac{\partial h_{11}}{\partial f_{11}} \frac{\partial f_{11}}{\partial w_3} = \boldsymbol{w_{11} h_{11} (1 - h_{11})}$
- $\frac{\partial \hat{y}}{\partial w_4} = \frac{\partial \hat{y}}{\partial h_{12}} \frac{\partial h_{12}}{\partial f_{12}} \frac{\partial f_{12}}{\partial w_4} = \boldsymbol{w_{12} h_{12} (1 - h_{12}) x_1}$
- $\frac{\partial \hat{y}}{\partial w_5} = \frac{\partial \hat{y}}{\partial h_{12}} \frac{\partial h_{12}}{\partial f_{12}} \frac{\partial f_{12}}{\partial w_5} = \boldsymbol{w_{12} h_{12} (1 - h_{12}) x_2}$
- $\frac{\partial \hat{y}}{\partial w_6} = \frac{\partial \hat{y}}{\partial h_{12}} \frac{\partial h_{12}}{\partial f_{12}} \frac{\partial f_{12}}{\partial w_6} = \boldsymbol{w_{13} h_{13} (1 - h_{13})}$
- $\frac{\partial \hat{y}}{\partial w_7} = \frac{\partial \hat{y}}{\partial h_{13}} \frac{\partial h_{13}}{\partial f_{13}} \frac{\partial f_{13}}{\partial w_7} = \boldsymbol{w_{13} h_{13} (1 - h_{13}) x_1}$
- $\frac{\partial \hat{y}}{\partial w_8} = \frac{\partial \hat{y}}{\partial h_{13}} \frac{\partial h_{13}}{\partial f_{13}} \frac{\partial f_{13}}{\partial w_8} = \boldsymbol{w_{13} h_{13} (1 - h_{13}) x_2}$
- $\frac{\partial \hat{y}}{\partial w_9} = \frac{\partial \hat{y}}{\partial h_{13}} \frac{\partial h_{13}}{\partial f_{13}} \frac{\partial f_{13}}{\partial w_9} = \boldsymbol{w_{13} h_{13} (1 - h_{13})}$
- $\frac{\partial \hat{y}}{\partial w_{10}} = \boldsymbol{1}$
- $\frac{\partial \hat{y}}{\partial w_{11}} = \boldsymbol{h_{11}}$
- $\frac{\partial \hat{y}}{\partial w_{12}} = \boldsymbol{h_{12}}$
- $\frac{\partial \hat{y}}{\partial w_{13}} = \boldsymbol{h_{13}}$

En fin, la mise à jour de chaque paramétré est donnée par la formule :

$$\Delta \boldsymbol{w_j} = \boldsymbol{\alpha (y - \hat{y}) \frac{\partial \hat{y}}{\partial w_j}}$$

<u>Exercice 3 :</u> Application numérique sur les réseaux de neurones (4 points)

Soit le même réseau de neurones multicouches de l'exercice 2 décrit par le graphe suivant :



Soit la donnée x = (2, -1), y = 1 et soient les valeurs initiales des paramètres w définies comme suit :

$w_1 = 1$, $w_2 = 0.5$, $w_3 = -0.25$, $w_4 = 0.75$, $w_5 = 1$, $w_6 = 0.25$, $w_7 = 0.5$, $w_8 = 0.5$, $w_9 = -0.5$, $w_{10} = 1$, $w_{11} = -1$, $w_{12} = 0.5$, $w_{13} = 0.25$.

1) Calculer les sorties intermédiaires $\boldsymbol{f_{11}}$, $\boldsymbol{f_{12}}$, $\boldsymbol{f_{13}}$, $\boldsymbol{h_{11}}$, $\boldsymbol{h_{12}}$, $\boldsymbol{h_{13}}$, $\boldsymbol{f_{21}}$ ainsi que la sortie finale $\hat{y}$.

Solution :

f_{11} = 1.2500** , **f_{12} =0.7500** , **f_{13} = 0**, **h_{11} = 0.7773** , **h_{12} = 0.6792**, **h_{13} = 0.5000** , **f_{21} = 0.6873, $\hat{y}$ = 0.6873.

2) Calculer les mise à jour $\Delta \boldsymbol{w_j}$ ainsi que les paramètres $\boldsymbol{w_j}$ pour j = 1,…,13 après une itération de mise à jour (en considérant le paramètre d'apprentissage $\boldsymbol{\alpha} = \mathbf{0.1}$).

NB : La précision des calculs numériques est fixée à 4 chiffres après la virgule.

Solution :

Les mises à jour Δw_j :

Δw_1	Δw_2	Δw_3	Δw_4	Δw_5	Δw_6	Δw_7	Δw_8	Δw_9	Δw_{10}	Δw_{11}	Δw_{12}	Δw_{13}
-0,0108	0,0054	-0,0054	0,0068	-0,0034	0,0034	0,0039	-0,0020	0,0020	0,0313	0,0243	0,0212	0,0156

Les paramètres w_j pour j = 1,…,13

w1	w2	w3	w4	w5	w6	w7	w8	w9	w10	w11	w12	w13
0,9892	0,5054	-0,2554	0,7568	0,9966	0,2534	0,5039	0,4980	-0,4980	1,0313	-0,9757	0,5212	0,2656