Centre Universitaire de Mila
Institut: ST, Département: MI
Master 1 STIC

Examen final en « Apprentissage Automatique »
01 : 30h / Mila le : 18-05-2017
Documents non autorisés

Corrigé type de l'examen final

Exercice 1 (10 points).

Répondre avec vrai ou faux. Justifier votre réponse.

1- Une hypothèse qui fait une bonne prédiction

bien pour le reste des données.

Réponse : faux.

La généralisation est très importante pour les techniques On peut trouver une hypothèse qui fait une bonne prédiction ne se généralise pas bien pour le reste des données.

2- Le modèle $y = xe^{\alpha_1} + \alpha_2$ peut être entrainé en utilisant la régression linéaire

Réponse : vrai.

y est une fonction linéaire de x.

3- Le modèle $y = \alpha_1 e^{x} + \alpha_2$ peut être entrainé en utilisant la régression linéaire.

Réponse : faux.

y est une fonction non linéaire de x.

4- Le modèle $y = \log(z^{\alpha_1} e^{x}) + \alpha_2$ peut être entrainé en utilisant la régression linéaire.

Réponse : vrai.

On a $y = \log(z^{\alpha_1} e^{x}) + \alpha_2 = \log(z^{\alpha_1}) \log(e^{x}) + \alpha_2 = x\alpha_1 \log(z) + \alpha_2$

Donc y est une fonction linéaire de x.

5- Le résultat de la classification utilisant le classificateur de Bayes est influencé par les poids des classes (justifier avec la formule).

Réponse : vrai.

$$p(y = C_k | \mathbf{x} = x, \mathcal{D}) = \frac{p(\mathbf{x} = x | y = C_k, \mathcal{D}) p(y = C_k | \mathcal{D})}{\sum_{j=1}^{K} p(\mathbf{x} = x | y = C_j, \mathcal{D}) p(y = C_j | \mathcal{D})}$$

6- Supposons un point correctement classé et qui se situe loin de la frontière de décision. La frontière de décision dans le cas de SVM ou la régression logistique serait affectée par ce point.

Réponse : faux.

0 à ces points.

7- L'astuce du noyau permet de résoudre des SVM avec des espaces de descripteurs de grandes dimensions.

Réponse : vrai.

Dans la formulation du SVM, le calcul des descripteurs apparaissent sous forme de produits scalaires qui peuvent être calculés de manière efficace par des noyaux.

8- La validation croisée est une mauvaise approche pour valider un classificateur.

Réponse : faux.

La validation croisée est la meilleure méthode pour la validation des classificateurs.

9- En classification SVM, l coordonnées et la frontière de décision est calculée par la formule $-\frac{w_0}{\|w\|}$

Réponse : vrai.

(Voir le support du cours pour la preuve.)

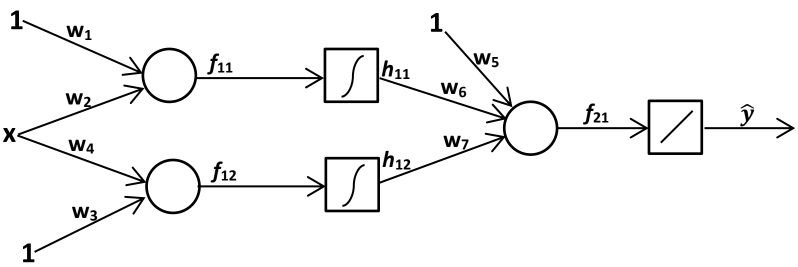
10- La méthode de montée du gradient est utilisée pour la régularisation des approches neurones.

Réponse : faux.

les optimum locaux (maxima ou minima).

<u>Exercice 2 (10 points).</u>

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



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

2- Soit

$$\mathrm{E(w)} = \left(y - \hat{y}_{(x,w)}\right)^2$$

En propagation en arrière (backpropagation), trouver les expressions des mises à jour des paramètres Δw_j pour j = 1, , 7.

Solution :

1- Propagation en avant (forward propagartion)

$$f_{11} = w_2 x + w_1$$

$$f_{12} = w_4 x + w_3$$

$$h_{11} = sigm(f_{11}) = \frac{1}{1 + e^{-f_{11}}}$$

$$h_{12} = sigm(f_{12}) = \frac{1}{1 + e^{-f_{12}}}$$

$$\hat{y} = \boldsymbol{f_{21}} = \boldsymbol{w_6 h_{11}} + \boldsymbol{w_7 h_{12}} + \boldsymbol{w_5}$$

2- Propagation en arrière (backpropagartion algorithm)

$$\mathrm{E(w)} = \left(y - \hat{y}_{(x,w)}\right)^2$$

Donc, on aura

$$\frac{\partial E(w)}{\partial w_j} = -2\left(y - \hat{y}_{(x,w)}\right)\frac{\partial \hat{y}_{(x,w)}}{\partial w_j}$$

:

$$\hat{y} = f_{21} = w_6 h_{11} + w_7 h_{12} + w_5$$

Donc, les dérivées $\frac{\partial \hat{y}_{(x,w)}}{\partial w_j}$ peuvent être calculées par :

$$\frac{\boldsymbol{\partial \hat{y}}}{\boldsymbol{\partial w_5}} = 1$$

$$\frac{\boldsymbol{\partial \hat{y}}}{\boldsymbol{\partial w_6}} = h_{11}$$

$$\frac{\boldsymbol{\partial \hat{y}}}{\boldsymbol{\partial w_7}} = h_{12}$$

$$\frac{\boldsymbol{\partial \hat{y}}}{\boldsymbol{\partial w_1}} = \frac{\partial \hat{y}}{\partial h_{11}}\frac{\partial h_{11}}{\partial f_{11}}\frac{\partial f_{11}}{\partial w_1} = w_6 h_{11}(1 - h_{11})$$

$$\frac{\boldsymbol{\partial \hat{y}}}{\boldsymbol{\partial w_2}} = \frac{\partial \hat{y}}{\partial h_{11}}\frac{\partial h_{11}}{\partial f_{11}}\frac{\partial f_{11}}{\partial w_2} = w_6 h_{11}(1 - h_{11})x$$

$$\frac{\boldsymbol{\partial \hat{y}}}{\boldsymbol{\partial w_3}} = \frac{\partial \hat{y}}{\partial h_{12}}\frac{\partial h_{12}}{\partial f_{12}}\frac{\partial f_{12}}{\partial w_3} = w_7 h_{12}(1 - h_{12})$$

$$\frac{\boldsymbol{\partial \hat{y}}}{\boldsymbol{\partial w_4}} = \frac{\partial \hat{y}}{\partial h_{12}}\frac{\partial h_{12}}{\partial f_{12}}\frac{\partial f_{12}}{\partial w_4} = w_7 h_{12}(1 - h_{12})x$$

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

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