Centre Universitaire de Mila

Institut ST
Département MI

Master STIC – S2
Mila le 23-05-2017

Corrigé type de l'interrogation 2 en apprentissage automatique

Exercice 01 :

Montrer que la fonction sigmoïde satisfait la propriété suivante :

$$\mathbf{sigm(-a) = 1 - sigm(a)}$$

et que son inverse est donné par

$$\mathbf{sigm^{-1}(y) = ln\{y/(1-y)\}}$$

Solution :

Soit

$$\mathbf{sigm(a) = \frac{1}{1+e^{-a}}}$$

Donc on a

$$\mathbf{1 - sigm(a) = 1 - \frac{1}{1+e^{-a}} = \frac{1+e^{-a}-1}{1+e^{-a}}}$$

$$\mathbf{= \frac{e^{-a}}{1+e^{-a}} = \frac{e^{a}e^{-a}}{e^{a}(1+e^{-a})} = \frac{1}{1+e^{a}} = sigm(-a)}$$

L'inverse de la fonction sigmoid peut être facilement trouvé comme suit

$$\mathbf{y = sigm(a) = \frac{1}{1+e^{-a}}}$$

$$\mathbf{\Rightarrow \frac{1}{y} - 1 = e^{-a}}$$

$$\mathbf{\Rightarrow ln\left\{\frac{1}{y} - 1\right\} = -a}$$

$$\mathbf{\Rightarrow ln\{y/(1-y)\} = a = sigm^{-1}(y)}$$

Exercice 02 :

Soit la fonction d'erreur pour une classification binaire dérivée pour un réseau de neurones ayant une fonction d'activation sigmoïde suivante :

$$\boldsymbol{E}(\widetilde{\mathbf{w}}) = \sum_{i=1}^{N} -\boldsymbol{y}^{(i)} \ln\left(\boldsymbol{g}\left(\widetilde{\boldsymbol{x}}^{(i)}\right)\right) - \left(\mathbf{1} - \boldsymbol{y}^{(i)}\right) \ln\left(\mathbf{1} - \boldsymbol{g}\left(\widetilde{\boldsymbol{x}}^{(i)}\right)\right)$$

Où $0 \leq g\left(\tilde{x}^{(i)}\right) \leq 1$, et les étiquettes des données $y^{(i)} \in \{0,1\}$.

Dériver la fonction d'erreur correspondante si on considère un réseau ayant $-1 \leq g\left(\tilde{x}^{(i)}\right) \leq 1$ et les valeurs des étiquettes $y^{(i)} = 1$ pour la classe C_1 et $y^{(i)} = -1$ pour la classe C_2. Quel est le meilleure choix pour la fonction d'activation de l'unité de sortie ?

Solution :

Cette transformation correspond simplement à un échelonnement et un décalage des sorties binaires, ce qui donne directement la fonction d'activation, sous la forme

$$\boldsymbol{y} = \mathbf{2}\boldsymbol{sigm}(\boldsymbol{x}) - \mathbf{1}$$

La fonction d'erreur correspondante peut être construite en appliquant la fonction inverse à $\boldsymbol{g}\left(\widetilde{\boldsymbol{x}}^{(i)}\right)$ et $\boldsymbol{y}^{(i)}$, ce qui va donner

$$\boldsymbol{E}(\widetilde{\mathbf{w}}) = \sum_{i=1}^{N} -\frac{\mathbf{1} + \boldsymbol{y}^{(i)}}{\mathbf{2}} \ln\left(\frac{\mathbf{1} + \boldsymbol{g}\left(\widetilde{\boldsymbol{x}}^{(i)}\right)}{\mathbf{2}}\right) - \left(\mathbf{1} - \frac{\mathbf{1} + \boldsymbol{y}^{(i)}}{\mathbf{2}}\right) \ln\left(\mathbf{1} - \frac{\mathbf{1} + \boldsymbol{g}\left(\widetilde{\boldsymbol{x}}^{(i)}\right)}{\mathbf{2}}\right)$$

$$\boldsymbol{E}(\widetilde{\mathbf{w}}) = \frac{\mathbf{1}}{\mathbf{2}} \sum_{i=1}^{N} \left\{-\left(\mathbf{1} + \boldsymbol{y}^{(i)}\right) \ln\left(\mathbf{1} + \boldsymbol{g}\left(\widetilde{\boldsymbol{x}}^{(i)}\right)\right) - \left(\mathbf{1} - \boldsymbol{y}^{(i)}\right) \ln\left(\mathbf{1} - \boldsymbol{g}\left(\widetilde{\boldsymbol{x}}^{(i)}\right)\right)\right\} + \boldsymbol{Nln}(\mathbf{2})$$

où le dernier terme peut être abandonné, car il est indépendant de w.

Pour trouver la fonction d'activation correspondante, nous appliquons simplement une transformation linéaire à la fonction sigmoïde, ce qui donne

$$\boldsymbol{g}(\boldsymbol{x}) = \mathbf{2}\boldsymbol{sigm}(\boldsymbol{x}) - \mathbf{1} = \frac{\mathbf{2}}{\mathbf{1} + \boldsymbol{e}^{-\boldsymbol{x}}} - \mathbf{1}$$

$$= \frac{\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{x}}}{\mathbf{1} + \boldsymbol{e}^{-\boldsymbol{x}}} = \frac{\boldsymbol{e}^{x/2} - \boldsymbol{e}^{-x/2}}{\boldsymbol{e}^{x/2} + \boldsymbol{e}^{-x/2}}$$

$$= \boldsymbol{tanh}(\boldsymbol{x}/\mathbf{2})$$