UNIVERSITÉ DE MONTRÉAL

Rapport : Devoir2

Pierre Gérard
Mathieu Bouchard

IFT3395-6390 Fondements de l'apprentissage machine
Pascal Vincent, Alexandre de Brébisson et César Laurent

Année académique 2015 - 2016

Table des matières

1 Partie théorique : Calcul du gradient pour l'optimisation des paramètres d'un réseau de neurones

1.1 Exercice a)

b est de dimension d_h
Le vecteur d'activation est : $h_a = W^{(1)}x + b$
Avec $h_{a_i} = W^{(1)}_{i1}x_1 W^{(1)}_{i2}x_2 + ... + W^{(1)}_{id}x_d b_i$
Et $h_{s_i} = h_{a_i} * I_{\{h_{a_i}>0\}} = max(h_{a_i}, 0)$

1.2 Exercice b)

$W^{(2)}$ est de dimension mxd_h
$b^{(2)}$ est de dimension m
Le vecteur d'activation est : $o^a = W^{(2)}h_s + b^{(2)}$
Avec $o^a_k = W^{(2)}_{k1}h_{s_1} + W^{(2)}_{k2}h_{s_2} + ... + W^{(2)}_{kn}h_{s_n} + b^{(2)}_k$

1.3 Exercice c)

$o^s = softmax(o^a) = \frac{1}{\sum_{i=1}^n e^{o_i^a}}(e^{o_1^a}, e^{o_2^a}, ..., e^{o_n^a})$
Donc $o^s_k = \frac{e^{o_k^a}}{\sum_{i=1}^n e^{o_i^a}}$
$e^x : \mathbb{R} \to \mathbb{R}^+$ donc la somme au numérateur de la fonction ci-dessus sera positive et la somme au numérateur aussi. Une fraction de deux nombres positifs sera toujours positif donc o^s_k est toujours positif.
$\sum_{i=1}^n o^s_i = \sum_{i=1}^n \frac{e^{o_i^a}}{\sum_{j=1}^n e^{o_j^a}}$
$= \frac{1}{\sum_{j=1}^n e^{o_j^a}} \sum_{i=1}^n e^{o_i^a}$
$= \frac{\sum_{i=1}^n e^{o_i^a}}{\sum_{j=1}^n e^{o_j^a}}$
$= 1$

C'est important car cela signifie que les sorties sont les probabilité pour l'entrée d'être une classe et ces classes sont mutuellement exclusives.

verifier l'explication

1.4 Exercice d)

$L(x, y) = -log o^s_y(x)$
$= -log \frac{e^{o_y^a(x)}}{\sum_{i=1}^n e^{o_i^a(x)}}$
$= -log(e^{o_y^a(x)}) + log(\sum_{i=1}^n e^{o_o^a(x)})$
$= -o^a_y(x) + log(\sum_{i=1}^n e^{o_i^a(x)})$