

Your Paper

You

January 1, 2017

1 Objectif

Le but principal de ce tp est d'implémenter un réseau de neurones multi-couches afin de classifier les images de la base MNIST.

2 Etapes

Les étapes de ce tp sont:

1. Implémenter la fonction sigmoïde et sa dérivée pour l'activation (somme pondérée des entrées de chaque couche sauf la dernière).

$$f(x) = \frac{1}{1 + \exp -x} \quad (1)$$

code:

```
limit = 30
z = np.clip(z, -limit, limit)
y = 1 / (1 + np.exp(-z))
yp = y * (1 - y)
return y, yp
```

2. Implémenter la fonction forward

$$y_1 = \text{activation}(w_0 \cdot w_0^t + b_0) \quad yp_1 = \text{la dérivée de } y \quad (2)$$

code:

```
for i in range(len(W)-1):
    F, Fp = act_func(W[i].dot(Y[i]) + B[i])
    Y.append(F)
    Yp.append(Fp)

Y.append(W[-1].dot(Y[-1]) + B[-1])

return Y, Yp
```

3. Implémenter la fonction backward

$$\text{grad}_{layer} = (W_k^t \cdot \text{grad} B_k) \text{grad} B_{k-1} = \text{grad}_{layer} * Y_{p_{k-1}} \quad (3)$$

code:

```

gradB = [error]
for i in reversed (range(1, len(W))):
    #print i
    gradB.insert(0,np.multiply(np.dot(W[i].T,gradB[0]),Yp[i-1]))
return gradB

```

4. Implementer la fonction update

$$w_{new} = w - \alpha \cdot \text{delta}_w \quad (4)$$

code:

```

for k in range(len(W)):
    grad_w = np.dot(gradB[k],Y[k].T)/batch_size
    grad_b = np.sum(gradB[k], axis=1)/batch_size
    W[k] = updateParams(W[k],grad_w, eta)
    B[k] = updateParams(B[k],grad_b.reshape(-1,1), eta)
grad_b.shape

```

3 Résultats

On note que le $n_epoch = 100$ et le $batch_size = 500$:

Paramètres	train accuracy	20% validation accuracy
eta = 1 et arch [100]	0.99782	0.9778
eta = 1 et arch [100,100]	0.9997	0.9789
eta = 0.1 et arch [100]	0.94772	0.9531
eta = 0.1 et arch [100,100]	0.94974	0.9537

On remarque que plus le nombre des couches est élevé les résultats sont meilleur également quand le pas est élevé on converge vite.

4 Difficultés rencontrés

La première difficulté réside dans la manipulation des matrices et listes mais j'en ai pu surmonter grâce à la documentation de numpy. Egalement la fonction update j'ai trouver des problèmes dans l'implémentation car je ne savais pas que j'aurais besoin de la fonction updateParams().