

Compte-rendu de TP2

CORENTIN LELOUP-THÉO DESCHAMPS-BERGER

21 novembre 2019

1 Données

Le Dataset est la base MNIST, composée de 50000 images de chiffres en résolution 28*28

2 Description de la tâche

Le but du TP est de construire une classe permettant de manipuler facilement les réseaux de neurones, d'en proposer des architectures différentes et de les entraîner.

On introduit pour cela une classe `tenseur` qui aura pour rôle de stocker les informations relatives à chaque couche. Chaque couche est représentée par un tenseur qui stocke toutes les informations relatives à la couche : données, gradient, pointeur vers la couche précédente etc .

Pour utiliser une telle classe, on a besoin d'implémenter les fonctions d'activation ainsi que leur dérivées pour passer le gradient lors de la backpropagation. On implémente ainsi *tanh* et *relu*. Il faut aussi implémenter les transformations affines ainsi que la fonction de perte et leurs backpropagations

On implémente également des fonctions capables d'initialiser les paramètres d'un réseau de neurone afin d'améliorer son optimisation, en l'occurrence les initialisations de Glorot et de Kaiming.

3 Evaluation du modèle

Grâce à notre nouvelle classe, on peut évaluer facilement différentes architectures et notamment des architectures multicouches sur notre dataset.

On observe une nette amélioration des résultats avec l'ajout d'une couche intermédiaire : on passe de 91,5% à 97%

Le choix de la fonction d'activation a une influence très limitée, on obtient les mêmes résultats en utilisant l'activation par tangente hyperbolique qu'avec la relu.