

Apprentissage statistique

TP3 : Réseaux de neurones

Olivier Schwander <olivier.schwander@lip6.fr>

2020-2021

Toute l'implémentation est à faire en PyTorch. Le but est de construire une architecture complète.

Question 1 *Exemple*

Téléchargez le fichier <https://github.com/pytorch/examples/blob/master/mnist/main.py>

Lancez ce fichier et analysez la sortie, puis le contenu du fichier.

Exercice 1 *Réseaux denses*

Question 1 *Perceptron*

Construisez un perceptron à plusieurs sorties.

Quelle fonction calcule ce perceptron ?

Question 2 *Couches cachées*

Empilez plusieurs couches denses.

Quelle fonction calcule ce réseau ? Y a-t-il un intérêt à ces couches cachées ?

Question 3 *Non-linéarités*

Pour rendre intéressantes les couches cachées, on a besoin d'introduire des non-linéarités dans l'empilement.

Les plus courantes sont :

- la tangente hyperbolique,
- la sigmoïde $S(x) = \frac{1}{1+e^{-x}}$,
- le rectifieur $f(x) = \max(0, x)$ (en anglais, *Rectified Linear Unit*, *ReLU*), c'est cette fonction qui est la plus utilisée en pratique.

Rajoutez des non-linéarités.

Question 4 *Soft-max*

Pour réaliser un classifieur multi-classes à partir d'un réseau à K sorties z_1, \dots, z_K , on utilise en général la fonction *soft-max* :

$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}$$

Quel est l'intérêt de cette fonction par rapport un simple maximum ?

Question 5

On utilisera comme pénalité l'entropie croisée :

$$L = \sum_{j=1}^N \sum_{i=1}^K t_j^{(i)} \log z_j^{(i)}$$

où $t_j = (0, \dots, 0, \underbrace{1}_k, 0, \dots, 0)$ si l'observation j appartient à la classe k et où z_j est la sortie du soft-max.

Ajoutez une fonction de loss à votre modèle.

Question 6

Expérimentez et comparez différents modèles à couches cachées pour les données USPS.

On pourra utiliser également la base d'image MNIST (accessible directement dans PyTorch, cf <https://pytorch.org/docs/stable/torchvision/datasets.html?#torchvision.datasets.MNIST>).

Question 7

Tracez les courbes d'apprentissage (accuracy et loss, en train et en validation, en fonction de l'époque). Commentez ces courbes.

Exercice 2 Réseaux convolutionnels**Question 1**

Quel est le nombre de paramètres d'un réseau dense à plusieurs couches cachées ?

Question 2

Quel est le nombre de paramètres pour un réseau convolutionnel avec le même nombre de couches ? Commentez.

Question 3

Construisez un réseau convolutionnel pour la classification MNIST.

Question 4

Expérimentez et analysez ce modèle.