

Exercice 1**Deep learning****4 points**

Quelle est la différence entre un CNN et un MLP complètement connecté (fully connected) ?

Exercice 2**Reconnaissance des voyelles****16 points**

Le but de cet exercice est de réaliser un système de reconnaissance de voyelles de l'anglais britannique par apprentissage à partir de données réelles. Le problème est spécifié par le fichier de données d'accompagnement, "VowelApp.mat" qui contient deux variables matlab :

- x un tableau $n \times p$ de $n = 528$ observations de $p = 10$ caractéristiques,
- S un vecteur de n étiquette indiquant la classe de l'objet (de 1 à 11)

Les étiquettes désignent les voyelles illustrés par le tableau suivant.

vowel	word	vowel	word
i	heed	O	hod
I	hid	C:	hoard
E	head	U	hood
A	had	u:	who'd
a:	hard	3:	heard
Y	hud		

1. un élève a résolu le problème en proposant la solution suivante en matlab

```
load VowelApp
[Xa, mu, sigma] = zscore(x);

y_b = eye(11);
for ii=1:length(S)
ya(ii,:) = y_b(S(ii),:);
end

rand('state',0)

[na p] = size(Xa);
nn = nnsetup([p 20 10 11]);

nn.learningRate = 1;
opts.numepochs = 5000;
opts.batchsize = na/8;
[nn, L] = nntrain(nn, Xa, ya, opts);

[er, bad] = nntest(nn, Xa, ya);

fprintf('\n resultat : %6.2f%% d''erreur \n', [er*100]);

% avec ce reseau de neurones vous aurez de l'ordre de 5,8 % d'erreur
```

Que pensez vous de ce programme et des performances annoncées ?

En python, on peut récupérer les données à l'aide des instructions suivantes :

```
import numpy as np
Xa=np.genfromtxt("Xa.txt", unpack=True)
S=np.genfromtxt("S.txt", unpack=True)
```

2. Il est possible de simplifier ce problème en le traitant comme un problème à deux classes : A vs. B
 - la classe A regroupe les voyelles étiquetés de 1 à 6 (le début du fichier)
 - la classe B regroupe les voyelles étiquetés de 7 à 11 (la fin du fichier)
Proposez un programme permettant de résoudre ce sous problème de classification à 2 classes
3. Proposez une méthode permettant d'apprendre un dictionnaire parcimonieux adapté à ce problème de classification à 2 classes
4. Représentez et visualisez en deux dimensions le tableau de données Xa