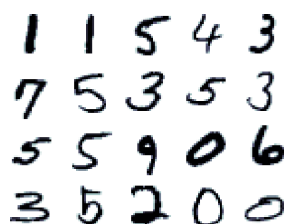


Reconnaissance de chiffres

Janvier 2016



L'objectif du TP est de développer un système de reconnaissance de chiffres manuscrits. Nous utiliserons le corpus MNIST.¹

1 Manipulation du corpus

Vous pouvez télécharger les données à partir du site web du cours. Le code suivant permet de charger les données :

```
1 import pickle
2 import gzip
3 import numpy
4
5 # Load the dataset
6 with gzip.open('mnist.pkl.gz', 'rb') as ifile :
7     train_set, valid_set, test_set = pickle.load(ifile, encoding='latin-1')
```

On récupère alors trois jeux de données, un pour l'apprentissage, un pour la validation et un pour le test. Ces jeux de données sont constitués d'un tableau regroupant deux objets :

- un tableau d'images, chaque image est elle-même un tableau de 784 valeurs (les pixels en niveau de gris entre 0 et 1) ;
- un tableau d'entiers correspondant aux chiffres à reconnaître (le i^e élément de ce tableau correspond à l'étiquette du i^e élément du tableau précédent).

Il est possible de visualiser une image en utilisant le code suivant :

1. <http://yann.lecun.com/exdb/mnist/>

```
1 import matplotlib.pyplot as plt
2
3 im = train_set[0][0].reshape(28,28)
4
5 plt.imshow(im, plt.cm.gray)
6 plt.show()
```

2 Prise en main

1. Quelle est la taille des images ? Comment sont-elles représentées dans les corpus ?
2. Comme premier traitement d'image, nous allons binariser l'image. Pour cela, il faut choisir un seuil sur le niveau de gris puis remplacer la valeur entière par 1 si le niveau de gris est supérieur à ce seuil, et 0 sinon. Programmer la binarisation de l'image. On utilisera un seuil à 0,5.²
3. Un histogramme horizontal est un tableau de même taille que le nombre de lignes de l'image. Ce tableau contient dans la i^{e} case, le nombre de pixel actif dans la i^{e} ligne de l'image. Écrire une fonction qui prend en paramètre une image et qui retourne son histogramme horizontal.
4. Afficher les histogrammes horizontal d'une image représentant un 1, d'une image représentant un 3 et une image représentant un 5. Pensez-vous que les histogrammes horizontaux fournissent une information pertinente pour la reconnaissance des chiffres ?
5. Par analogie avec la question précédente, écrivez une fonction qui retourne l'histogramme vertical d'une image.
6. Écrivez une fonction qui prend en paramètre une liste d'images et qui renvoie l'image moyenne correspondante. Dans cette image, la valeur du pixel à la position (x, y) est obtenue en calculant la moyenne des valeurs des pixels à la position (x, y) de l'ensemble des images passées en paramètres.
7. Représentez et commentez les images moyennes obtenues en considérant tous les exemples d'une classe du corpus d'apprentissage (c.-à-d. l'image moyenne de toutes les images représentant un 0, de toutes les images représentant un 1, ...).

3 Classification

On se propose de comparer les performances obtenues par deux méthodes de classification en utilisant plusieurs types de caractéristiques.

2. Pour les plus curieux, il est intéressant de voir ce qu'il se passe lorsque l'on considère plus bas — 0,4 par exemple.

8. Rappelez (en deux phrases) le principe des k plus proches voisins et du classifieur à distance minimum.
9. Quelle est la complexité de ces deux approches (en apprentissage et en test) ? Commentez.

On ne considère dans la suite que la représentation des chiffres sous forme d'images binarisées et la norme euclidienne comme distance.

10. Représentez l'erreur en classification obtenue sur les trois corpus en appliquant la méthode des k -ppv pour k variant dans $\llbracket 1, 10 \rrbracket$.
11. Comment choisir la valeur optimale de k . Quelle performance obtient-on alors ?
12. Implémentez un classifieur à distance minimum. Quel est sa performance en test et en apprentissage.
13. Déterminez la matrice de confusion de la méthode des k -ppv. La matrice de confusion est une matrice $n \times n$ où n est le nombre de classe. L'élément i, j de cette matrice correspond au nombre de fois où la i^{e} classe a été prédite alors que la vraie classe était la j^{e} . Commentez.
14. Représentez graphiquement 2 exemples mal classés. Commentez.

En ne considérant que la valeur optimale de k , déterminez maintenant les performances obtenues en :

15. considérant les images brutes (c.-à-d. non binarisée) ;
16. considérant uniquement les histogrammes verticaux et horizontaux des images.