

Apprentissage supervisé

Maxence Hanin, Louis Jean

20 Novembre 2019

1 Manipulation de la base de données

1.1 Exécution et description des commandes

Voici les différentes commandes exécutées ainsi que leur description :

```
print(mnist)
```

Cette commande donne les explications concernant le jeu de données, ainsi que sa structure.

```
print (mnist.data)
```

Cette commande donne l'ensemble des images sous forme d'une liste de 70000 sous-listes de 784 éléments, correspondant aux 784 pixels de l'image.

```
>>> print (mnist.data)
[[0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
 ...
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
```

```
print (mnist.target)
```

Cette commande donne la liste des 70000 valeurs (entre 0 et 9) que représentent chaque images.

```
len(mnist.data)
```

Cette commande renvoie 70000, soit la taille de data.

```
print (mnist.data.shape)
```

Cette commande renvoie (70000, 784) soit le couple (taille de data, taille de chaque élément).

```
print (mnist.target.shape)
```

Cette commande renvoie (70000,) soit le couple (taille du dataset, taille de chaque élément).

```
mnist.data[0]
```

Cette commande renvoie le premier élément de data, soit une liste de la valeur des 784 pixels .

```
mnist.data[0][1]
```

Cette commande renvoie la valeur du deuxième pixel de la première image de data.

```
mnist.data[:,1]
```

Cette commande renvoie un array contenant la colonne 2 de data (la liste des deuxièmes valeurs des pixels des 70000 images).

```
mnist.data[:100]
```

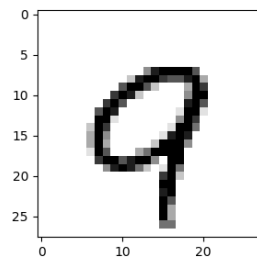
Cette commande renvoie les deux premières images

1.2 Visualisation des données

2 La méthode des k-nn

Exemple d'exécution de la méthode knn :

```
inge/TP1-knn$ python tp1.py --neighbors 3 --split 0.7 --distance 3 --multithread
train size = 3500, test size = 1500
Training with k = 3, n_jobs = -1, minkowski distance with p = 3.000000
Score for 3 neighbors and p = 3.000000 is 0.934000
Training time 363.216400 ms, scoring time 126501.708269 ms
['9']
```



```
inge/TP1-knn$ python tp1.py --visualize_data 5
2
```

