

Année 2020/2021

Master Informatique Visuelle

Intelligence Artificielle

Série de travaux pratiques 3

K proches voisins (KNN)

Exercice 1

Il s'agit d'appliquer l'algorithme K proches voisins (KNN) pour la classification des chiffres manuscrits. Pour ce faire, il faut :

1. Télécharger le data-set en utilisant la fonction :

```
sklearn.datasets.load_digits() # voir le lien
```

2. Utiliser l'attribut 'DESCR' pour afficher la description du data-set.
3. Quelle est la taille du data-set ?
4. Quel est le nombre de classes ?
5. Quel est le nombre d'attributs ?
6. Diviser le data-set en un ensemble d'apprentissage (70%) et un ensemble de test (30 %).
7. Créer un classifieur KNN de la bibliothèque Scikit-Learn :

```
neighbors.KNeighborsClassifier(k)
```

8. Par défaut, la métrique 'uniform' est utilisée pour calculer la distance. Comparer le résultat lorsque 'distance' est utilisée.
9. Classifier l'ensemble du data-set pour un K=3.
10. Calculer la précision, le rappel et le temps d'exécution.

```
from sklearn.metrics import recall_score, precision_score  
recall_score(y_true, y_pred)  
precision_score(y_true, y_pred)
```

Pour calculer le temps d'exécution, utiliser la fonction 'time' :

```
import time
start_time = time.time()
# Instructions
print("--- %s seconds ---" % (time.time() - start_time))
```

11. Varier le K (1, 3, 5, 7, 9, etc), comparer et visualiser en graphe les résultats obtenus.