

BUT Informatique 2^{ème} année Parcours A

R3.A.15: Machine Learning

2023-2024

Sébastien Lefèvre
sebastien.lefevre@univ-ubs.fr

Evaluation

- Correction en séance
- Discussion : modalité d'évaluation

R3.A.15

Évaluation 2

Durée : 10 minutes

2023-2024

Expliquer le principe d'un des algorithmes de classification supervisée vus lors de la séance précédente.

Nom de l'algorithme :

(Name in english :

)

Bilan séance 2

1. Découverte des principaux algorithmes de classification

2. Mise en œuvre sous scikit-learn

```
>>> from sklearn.datasets import load_iris
>>> from sklearn.model_selection import train_test_split
>>> from sklearn.naive_bayes import GaussianNB
>>> X, y = load_iris(return_X_y=True)
>>> X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.5, random_state=0)
>>> gnb = GaussianNB()
>>> y_pred = gnb.fit(X_train, y_train).predict(X_test)
>>> print("Number of mislabeled points out of a total %d points : %d"
... % (X_test.shape[0], (y_test != y_pred).sum()))
```

```
>>> from sklearn.datasets import load_iris
>>> from sklearn import tree
>>> iris = load_iris()
>>> X, y = iris.data, iris.target
>>> clf = tree.DecisionTreeClassifier()
>>> clf = clf.fit(X, y)
>>> tree.plot_tree(clf)
```

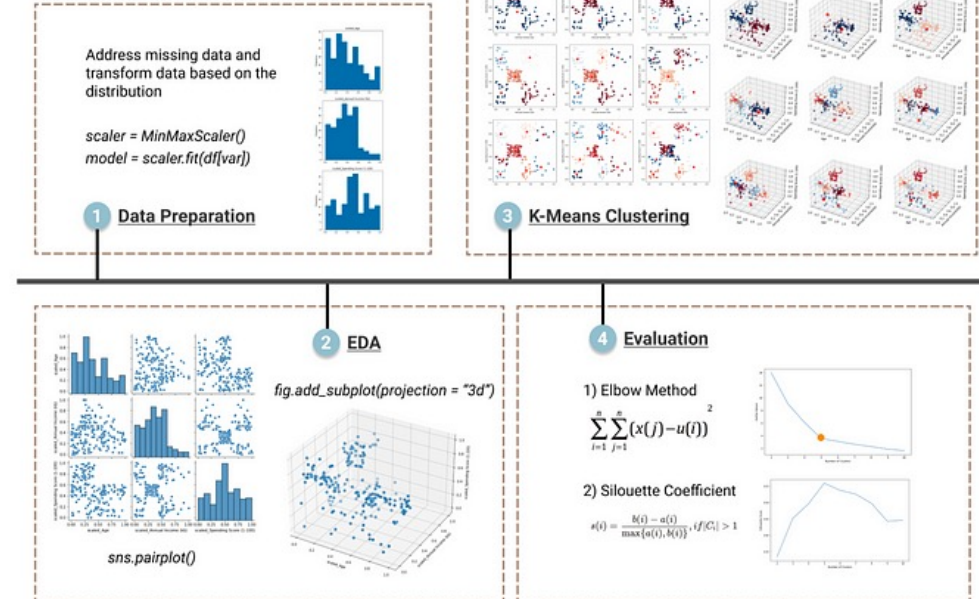
Séance 3 : Classification non-supervisée

1. Principes
2. Algorithmes
3. Mise en œuvre en Python

2.3. Clustering

- 2.3.1. Overview of clustering methods
- 2.3.2. K-means
- 2.3.3. Affinity Propagation
- 2.3.4. Mean Shift
- 2.3.5. Spectral clustering
- 2.3.6. Hierarchical clustering
- 2.3.7. DBSCAN
- 2.3.8. HDBSCAN
- 2.3.9. OPTICS
- 2.3.10. BIRCH
- 2.3.11. Clustering performance evaluation

CLUSTERING ANALYSIS & CUSTOMER SEGMENTATION.



Restitutions :

1. Explication d'un ou plusieurs algorithmes (fiche de révision)
2. Exemple(s) de code Python (notebook)