

TP N°2

Dans ce TP, nous appliquons un modèle de classification supervisée sur le célèbre dataset *Iris*. L'objectif est d'entraîner un Random Forest et d'évaluer sa capacité à prédire correctement l'espèce des fleurs.

Nous avons d'abord chargé les données et séparé les variables explicatives des étiquettes. Un découpage train/test à 80/20 a été réalisé, puis un modèle RandomForestClassifier a été entraîné sur les données d'apprentissage. Enfin, nous avons évalué les performances du modèle via l'accuracy, le classification report et la matrice de confusion.

```
PS C:\Users\sawer\Documents\M1IB2\tp_supervise\atelier_unsup\TP2>
```

```
Classification report:
      precision    recall  f1-score   support

      Setosa       1.00     1.00     1.00      10
  Versicolor     0.82     0.90     0.86      10
  Virginica     0.89     0.80     0.84      10

  accuracy           0.90      0.90      0.90      30
  macro avg       0.90     0.90     0.90      30
  weighted avg    0.90     0.90     0.90      30

  macro avg       0.90     0.90     0.90      30
  weighted avg    0.90     0.90     0.90      30
```

Matrice de confusion:

```
[[10  0  0]
 [ 0  9  1]
 [ 0  2  8]]
[[10  0  0]
 [ 0  9  1]
 [ 0  2  8]]
[ 0  9  1]
 [ 0  2  8]]
```

Le Random Forest obtient de très bonnes performances avec une accuracy de 90 %, confirmant sa robustesse sur ce type de données. La classe Setosa est parfaitement reconnue, tandis que quelques confusions apparaissent entre Versicolor et Virginica, ce qui est attendu. Globalement, les résultats montrent que le modèle est bien adapté au dataset Iris.