

Exercices

Chapitre 6 : Apprentissage supervisé - La classification

Hiver 2025

Étape 1 : Préparation des données

1. Téléchargez les données de l'ensemble de données iris
2. Déterminez le nombre de classe, d'individus par classe et de caractéristiques
3. Répartissez les données en deux ensembles égaux : `train` et `test` (`random_state = 10`). Vérifiez le nombre d'individus dans chaque classe de l'ensemble `train`
4. Représentez le diagramme de dispersion des paires de variables (utilisez `sns.pairplot`)
5. Répartissez les données en deux ensembles 2/3 pour `train` et 1/3 pour `test` (`random_state = 10`). Vérifiez le nombre d'individus dans chaque classe de l'ensemble `train`
6. Créer des matrices de caractéristiques (`x_train`, `x_test`), des vecteurs de classes (`y_train`, `y_test`)

Étape 2 : Classification

1. Construisez un arbre de décision (`hauteur = 3`, `random_state = 1`)
2. Visualisez l'arbre de décision
3. Testez l'arbre de décision sur les données de test
4. Représentez la matrice de confusion. Déterminez les taux de classification par classe. Commentez les résultats.
5. Refaire toutes les étapes précédente pour une hauteur = 2 et une hauteur = 4.