

Model-based learning

Travaux pratiques sous le logiciel R

TP1 : programmation de l'algorithme des kmeans

Dans ce premier TP, l'objectif sera de coder en R un algorithme simple de clustering.

1. Programmer l'algorithme des kmeans sous R. Vous utiliserez la distance euclidienne. La fonction devra prendre en entrée la matrice de données et le nombre de clusters souhaités. Une option permettra de fixer le nombre d'initialisations. En sortie, vous retournerez la partition ainsi que l'inertie intra-classes.
2. Tester votre algorithme sur les données iris et le comparer à l'algorithme kmeans de R (en terme de partition et inertie intra-classes).