

## TP3 M2 - INFO

### Exercice 1

- 1) Charger et analyser les packages MASS, klaR, mda.
- 2) Réaliser une ACP sur le jeu de données Iris en coloriant les individus selon leur classes. Sauvegardez le graphique sous forme d'un fichier de type .pdf.
- 3) Créer à partir de jeu de données Iris, par tirage aléatoire, un échantillon de données utilisé pour l'apprentissage (80% des données) et un échantillon de test.
- 4) Déterminer pour l'échantillon d'apprentissage l'effectif de chaque classe.
- 5) Renommer les classes du jeu de données Iris par : 'a' pour 'virginica', par 'e' pour 'versicolor' et par 's' pour 'species'.
- 6) Réaliser une description graphique des groupe à l'aide de la fonction *partimat()*.
- 7) Calculer un ADL à l'aide de la fonction *fda()*.
- 8) Calculer les observations mal classées.
- 9) Représenter les résultats de l'ADL sous forme d'un graphique. Comparer avec le résultat obtenu à l'aide de l'ACP. Qu'est-ce que vous observez? Réalisez le même scénario sur le jeu de données (*Waveform*).

### Exercice 2

- 1) Analysez la fonction *dist()* avec toutes les options proposées. Réaliser l'exemple présenté dans la documentation.
- 2) Analyser la fonction *cmdscale()* du package MASS. Appliquer cette fonction sur le jeu de données «eurodist».
- 3) Installer les package MASS, RDRToolbox, golubEsets. Analyser la documentation de ces packages.
- 4) Charger le jeu de données « GolubEsets». Analyser et faire une courte description de ce jeu de données.
- 5) Réaliser une ACP sur ce jeu de données. Représenter graphiquement les observations en utilisant leurs coordonnées sur le 1 et 2<sup>ème</sup> axe.
- 6) Réaliser une MDS sur ce jeu de données en utilisant la fonction *cmdscale()*. Qu'est-ce que vous observez par rapport aux résultats obtenus avec l'ACP ?
- 7) Analyser la fonction *LLE*. Appliquer cette fonction sur les jeux de données proposés en faisant varier le nombre de voisins par 3, 5, 8, 10, 12 et 15. Analyser les résultats et déterminer le meilleur paramètre de voisinage.
- 8) Analyser la fonction *Isomap*. Appliquer cette fonction sur les jeux de données proposés en faisant varier le nombre de voisins par 4, 8, 10, 13, 16 et 20. Analyser les résultats et déterminer le meilleur paramètre de voisinage.
- 9) Charger le jeu de données «Alon». Analyser et faire une courte description de ce jeu de données. Appliquer les approches vues dans les étapes précédentes et analyser les résultats.