

Programme de remise à niveau sur le thème « Data Mining – Analyse prédictive ».

1. Supports à lire

Les supports suivants présentent les fondements de l'apprentissage supervisé. Ils seront complétés par des documents spécifiques aux thèmes abordés, référencés dans les sujets des exercices à faire.

- Denis F., Gilleron R., « Apprentissage à partir d'exemples – Notes de cours » ;
<http://www.grappa.univ-lille3.fr/polys/apprentissage/>
- Rakotomalala R., « Apprentissage supervisé » ;
http://eric.univ-lyon2.fr/~ricco/cours/slides/Apprentissage_Supervise.pdf
<http://eric.univ-lyon2.fr/~ricco/tanagra/fichiers/643-2464-2-PB.pdf>
- Habermehl P., Kesner D., « Algorithmes d'apprentissage » ;
<https://www.irif.univ-paris-diderot.fr/~kesner/enseignement/iup/cours71.pdf>

2. Exercices à réaliser

Il est illusoire de vouloir tout comprendre avant de vous lancer dans les exercices sur machine. Après avoir lu rapidement les supports, essayez de commencer. Vous aurez alors le loisir de revenir plus spécifiquement sur les différents thèmes quand vous chercherez à répondre précisément aux questions.

Autre conseil important, concentrez-vous sur les aspects opérationnels. Il sera toujours temps par la suite de comprendre les principes fondateurs. A chaque problème à traiter doit correspondre un ou plusieurs fichiers à retourner (**Excel** ou **R** ou **Python**). Commentez autant que possible votre travail.

Problèmes à traiter :

1. SISE – DM 1 – Analyse discriminante prédictive (breast cancer)
2. SISE – DM 1 – Analyse discriminante et sélection de variables (seeds)
3. SISE – DM 1 – Arbres de décision, lecture des règles et arbres interactifs (heart)
4. SISE – DM 1 – Arbres de décision, frontières de décision (iris)
5. SISE – DM 1 – Comparaison des méthodes (segmentation)