

| <b>Données Intelligentes</b><br><i>Smart Data</i>          |  |  |  |
|--|--|--|--|
| <b>Code ECUE</b> <i>Course code:</i> <b>DOI</b>            |  |  | <b>UE : UE5-2d</b>                             |
| <b>Département</b> <i>Department</i>                       | IA   |  | <b>Cours</b> <i>Lectures</i> 15h00             |
| <b>Coordonnateurs</b> <i>Lecturers</i>                     | B. Chardin                                 |  | <b>T.D.</b> <i>Tutorials</i> 15h00             |
| <b>Période</b> <i>Year of study</i>                        | A3   |  | <b>T.P.</b> <i>Laboratory sessions</i>         |
| <b>Semestre</b> <i>Semester</i>                            | S5   |  | <b>Projet</b> <i>Project</i>                   |
| <b>Evaluation</b> <i>Assessment method(s)</i>              | 1 écrit                                    |  | <b>Non encadré</b> <i>Unsupervised</i>         |
| <b>Langue d'instruction</b> <i>Language of instruction</i> | Français                                   |  | <b>Horaire global</b> <i>Total hours</i> 30h00 |
| <b>Type de cours</b> <i>Type of course</i>                 | Option Données                             |  | <b>Travail personnel</b> <i>Homework</i> 10h00 |
| <b>Niveau</b> <i>Level of course</i>                       | Second cycle universitaire <i>Graduate</i> |  |  |

#### Compétences attendues :

- Modéliser un problème de décision en termes de programme d'optimisation linéaire.
- Savoir utiliser un solveur et interpréter les résultats.
- Comprendre l'importance du prétraitement des données et connaître les techniques applicables.
- Comprendre et savoir utiliser les techniques principales d'extraction de connaissances et d'apprentissage automatique

**Pré-requis :** Algorithmes et Systèmes numériques, Utilisation et exploitation des données, Méthodes numériques et Calcul scientifique, Apprentissage Automatique Appliqué

#### Contenu :

- Programmation par contrainte
  - Programmation linéaire
  - Programmation linéaire en nombre entiers
  - Problèmes de satisfaction de contraintes
  - Processus de modélisation, résolution et interprétation
  - Implémentation avec IBM ILOG CPLEX
- Prétraitement des données
  - Impact de la qualité des données sur leur traitement
  - Identification et correction des données erronées
  - Processus ETL (Extract, Transform, Load)
- Fouille de données
  - itemsets fréquents
  - Règles d'association
  - Dépendances fonctionnelles et variantes
  - Mesures d'intérêt applicables aux règles générées
- Apprentissage automatique
  - Problèmes de classification et de régression
  - Classification non supervisée (partitionnement)
  - Apprentissage supervisé (arbres et forêts de décision, SVM, réseaux de neurones)
  - Méthodologie de validation expérimentale

#### Bibliographie :

- Dantzig George, Thapa Mukund, Linear Programming 1: Introduction. Springer Series in Operations Research and Financial Engineering, 2003.

- Dantzig George, Thapa Mukund, Linear Programming 2: Theory and Extensions. Springer Series in Operations Research and Financial Engineering, 2003.
- Marriott , K., Stuckey, P., Programming with constraints : an introduction, The MIT Press, 1998
- Mazhar Hameed and Felix Naumann. 2020. Data Preparation: A Survey of Commercial Tools. SIGMOD Rec. 49, 3 (September 2020), 18–29.
- Panik, M.J., Linear Programming: Mathematics, Theory and Algorithms, 1996.
- Vasilyev, F.P., Ivanitskiy, A.Y., In-Depth Analysis of Linear Programming, 2001
- Wolsey, L., Integer programming, Wiley-Interscience series in discrete mathematics and optimization. J., Wiley & sons, New York (N.Y.), Chichester, Weinheim, 1998