

Analyse de données et data mining

| | |
|--|-----------------------------------|
| Course title – Intitulé du cours | Analyse de données et data mining |
| Level / Semester – Niveau /semestre | M2 Economie Appliquée |
| School – Composante | Ecole d'Economie de Toulouse |
| Teacher – Enseignant responsable | Julien ALIGON et Moncef GAROUANI |
| Other teacher(s) – Autre(s) enseignant(s) | |
| Other teacher(s) – Autre(s) enseignant(s) | |
| Other teacher(s) – Autre(s) enseignant(s) | |
| Other teacher(s) – Autre(s) enseignant(s) | |
| Other teacher(s) – Autre(s) enseignant(s) | |
| Lecture Hours – Volume Horaire CM | 30h |
| TA Hours – Volume horaire TD | |
| TP Hours – Volume horaire TP | |
| Course Language – Langue du cours | |
| TA and/or TP Language – Langue des TD et/ou TP | |

Teaching staff contacts – Coordonnées de l'équipe pédagogique :

ALIGON Julien ; julien.aligon@ut-capitole.fr ; MF229

GAROUANI Moncef ; moncef.garouani@ut-capitole.fr ; MF229

Course Objectives – Objectifs du cours :

Une première partie de ce cours a pour objectif d'introduire les fondamentaux en fouille de données (motifs fréquents) et en apprentissage automatique (apprentissage non-supervisé et supervisé). Il s'agit d'y donner les bonnes pratiques à l'utilisation de ces technologies.

La deuxième partie se concentrera sur l'analyse des résultats des prédictions dans le cadre de l'apprentissage supervisé, par l'utilisation de techniques d'explicabilité (XAI). En effet, les modèles supervisés sont souvent vus comme des boîtes noires, pour lesquels il est pourtant important d'en comprendre le fonctionnement (en particulier pour des non-experts). Une introduction à l'explicabilité sera ainsi proposée et un focus particulier sera donné aux méthodes SHAP, très utilisées de nos jours dans les domaines de la finance, médecine, etc.

Prerequisites – Pré requis :

- Pas de pré-requis

Practical information about the sessions – Modalités pratiques de gestion du cours :

Le cours sera surtout axé sur la pratique par une mise en œuvre des concepts du cours via l'utilisation de notebooks et de l'analyse de données réels, majoritairement issus de littérature scientifique.

Grading system – Modalités d'évaluation :

- Contrôle sur table (partie machine learning)
- Projet (partie XAI)

Bibliography/references – Bibliographie/références :

- John D. Kelleher et al., Fundamentals of Machine Learning for Predictive Data Analytics, second edition: Algorithms, Worked Examples, and Case Studies, ISBN : 978-0262044691
- Denis Rothman, Hands-On Explainable AI (XAI) with Python: Interpret, visualize, explain, and integrate reliable AI for fair, secure, and trustworthy AI apps. Packt Publishing; 1er édition (31 juillet 2020), ISBN: 1800208138.
- Christoph Molnar, Interpretable Machine Learning: A Guide for Making Black Box Models Explainable, ISBN : 979-8411463330

- D'autres ressources à partager en classe.

Session planning – Planification des séances

Les séances sont divisées en 10 séances de 3h chacune, selon le programme suivant :

- 1^{ère} séance : introduction à la programmation python
- 2^{ème} séance à 5^{ème} séance : introduction à l'intelligence artificielle et au machine learning
- 6^{ème} séance à 10^{ème} séance : analyse prédictive et XAI (eXplainable Artificial Intelligence)