

Apprentissage Statistique

Course title – Intitulé du cours	Apprentissage Statistique
Level / Semester – Niveau /semestre	M2/S2
School – Composante	Ecole d'Economie de Toulouse
Teacher – Enseignant responsable	Pascal Lavergne
Lecture Hours – Volume Horaire CM	30
TA Hours – Volume horaire TD	6
TP Hours – Volume horaire TP	
Course Language – Langue du cours	Francais
TA and/or TP Language – Langue des TD et/ou TP	

Teaching staff contacts – Coordonnées de l'équipe pédagogique :

Mail : pascal.lavergne@ut-capitole.fr Bureau T 506

Accueil des étudiants: sur rendez-vous

Course Objectives – Objectifs du cours :

Dans ce cours nous allons apprendre

- les principaux concepts d'apprentissage statistique : régression et classification
- les principales méthodes : (i) Régression linéaire (avec beaucoup de régresseurs) (ii) Régression non-linéaire (iii) Méthodes non paramétriques (iv) Arbres (v) Boosting (vi) Support Vector Machines
- comment appliquer et évaluer ces méthodes : validation croisée et bootstrap
- comment interpréter les résultats

Ces compétences sont utiles pour le marketing, l'analyse macroéconomique, l'évaluation des politiques publiques, la prévision, et l'analyse client, entre autres.

Prerequisites – Pré requis :

- Connaissances en analyse / algèbre (niveau M1)
- Probabilités et statistiques (niveau L3)
- Cours d'Econométrie et d'Econométrie Appliquée de M1 ou équivalents
- Connaissance de R et RStudio
- Connaissance de RMarkdown

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

Mercredi 14h-17h (à confirmer)

Grading system – Modalités d'évaluation :

L'évaluation des étudiants sera basée sur

1. La participation individuelle en cours
2. Des devoirs maisons rédigés en groupe
3. Un projet rédigé en groupe

Bibliography/references – Bibliographie/références :

- **Requis :** G. James, D. Witten, T. Hastie and R. Tibshirani (2017) An Introduction to Statistical Learning with Applications in R, 8th edition, Springer Texts in Statistics. Livre disponible gratuitement (et téléchargeable) via l'ENT
- The Caret package cf. <https://topepo.github.io/caret/>

- T. Hastie, J. Friedman and R. Tibshirani W. A. Woodward, H.L. Gray, and A.C. Elliot (2013) The Elements of Statistical Learning, Springer Texts in Statistics.
- M. Kuhn and K. Johnson (2013) Applied Predictive Modelling Springer Texts in Statistics.
- Wikistat Science des Données, Apprentissage Statistique & IA <https://github.com/wikistat/Apprentissage>

Session planning – Planification des séances

Sera communiquée ultérieurement aux étudiants.

Distance learning – Enseignement à distance :

En cas de nécessité, un enseignement à distance sera assuré en mobilisant, par exemple

- Classe en ligne interactive
- Quizzes et exercices en ligne
- Forums