

## **Online and Reinforcement Learning**

Course title – Intitulé du cours	Online and Reinforcement Learning
Level / Semester – Niveau /semestre	M2, Semester 2
School – Composante	Ecole d'Economie de Toulouse
Teacher – Enseignant responsable	Julien Chhor
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	15h
TA Hours – Volume horaire TD	
TP Hours – Volume horaire TP	
Course Language – Langue du cours	English
TA and/or TP Language – Langue des TD et/ou TP	

### **Teaching staff contacts – Coordonnées de l'équipe pédagogique :**

[julien.chhor@tse-fr.eu](mailto:julien.chhor@tse-fr.eu)

### **Course Objectives – Objectifs du cours :**

Online learning is a popular framework in statistics and machine learning where the data set is not entirely available at the beginning and is rather revealed sequentially (that is, one data point at a time). At each time step, the learner has to predict an outcome based on the currently available information, before the next data point is revealed. The goal is to develop a strategy that improves over time to make as accurate predictions as possible eventually. We will cover the following topics

- Online classification in realizable case, halving.
- Online gradient descent for convex and strongly convex loss. Online-to-batch conversion. Online linear regression.
- Randomization by exponential weighting. Prediction with expert advice.

The second part of the course focuses on reinforcement learning, an area of machine learning in which an agent learns how to behave in an environment by performing actions and assessing the results. We will cover the following topics

- Adversarial multi-armed bandit problem, exploration vs exploitation trade-off.
- Markov decision processes, Bellman equations, dynamic programming.
- Q-learning

**Prerequisites – Pré requis :**

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

**Grading system – Modalités d'évaluation :**

**Bibliography/references – Bibliographie/références :**

Shalev-Schwartz, S. (2011) *Online learning and online convex optimisation*. Foundations and Trends in Machine Learning, vol. 4.

Lattimore, T., & Szepesvári, C. (2020). *Bandit Algorithms*. Cambridge University Press.

**Session planning – Planification des séances**

**Distance learning – Enseignement à distance :**