

# 3MD3220 - Apprentissage par renforcement

Responsables : **Fragkiskos MALLIAROS**

Langues d'enseignement : **FRANCAIS**

Campus où le cours est proposé : **CAMPUS DE PARIS - SACLAY**

Nombre d'heures d'études élèves (HEE) : **40**

Nombre d'heures présentielles d'enseignement (HPE) : **24**

Année académique : **2024-2025**

Niveau avancé : **non**

---

## Présentation, objectifs généraux du cours :

Les méthodes d'apprentissage par renforcement (RL) sont à l'intersection de plusieurs domaines tels que l'apprentissage automatique, la théorie des probabilités, la théorie du contrôle ou la théorie des jeux. Elles étudient les compromis fondamentaux entre exploitation et exploration afin d'optimiser une récompense future dans un environnement inconnu.

Récemment, elles ont été appliquées avec succès à de nombreuses applications telles que des jeux complexes (Go ou Atari), le contrôle de robots, la sélection d'experts, conduisant dans de nombreux cas à des performances surhumaines dans ces contextes. Ce cours couvrira les outils nécessaires (d'un point de vue théorique et pratique) pour comprendre comment une telle percée a été possible.

## Période(s) du cours (n° de séquence ou hors séquence) :

SM11

## Prérequis :

Les étudiants doivent être familiarisés avec les concepts de base des probabilités, tels que la chaîne de Markov, ainsi qu'avec les fondements des concepts d'apprentissage automatique.

## Plan détaillé du cours (contenu) :

1. Introduction à RL
2. PDM et programmation dynamique
3. Bandits
4. Outils d'évaluation
5. Outils de contrôle

6. RL avec approximations

7. S'aventurer dans la RL profonde

## Déroulement, organisation du cours :

Chaque section du cours sera soit un cours magistral de 1h30, soit un laboratoire de 1h30. Les laboratoires comprendront des sessions pratiques et des petits projets en python et donneront aux étudiants l'occasion de tester les concepts théoriques sur des exemples concrets.

## Organisation de l'évaluation :

L'évaluation du cours sera basée sur les éléments suivants :

- Devoirs : les devoirs comprendront des questions théoriques ainsi que des codes pratiques qui devront s'exécuter sur des notebooks jupyter.

- Projet : Les étudiants doivent former des groupes de 2 à 3 personnes, proposer un sujet pour leur projet (qui devra être validé), et soumettre un rapport final de projet.

## Acquis d'apprentissage visés dans le cours :

Ce cours a pour but d'initier les étudiants à la compréhension des concepts sous-jacents des algorithmes d'apprentissage par renforcement. Au-delà de la théorie, les étudiants auront l'opportunité de pratiquer l'apprentissage par renforcement via des sessions de laboratoire.

## Bibliographie :

- Sutton & Barto, Reinforcement Learning: An Introduction
- Szepesvari, Algorithms for Reinforcement Learning
- Bertsekas, Dynamic Programming and Optimal Control, Vols I and II
- Puterman, Markov Decision Processes: Discrete Stochastic Dynamic Programming