

Apprentissage par renforcement et modèles génératifs I



Niveau
d'étude
BAC +5 /
master



ECTS
3 crédits



Composante
Faculté des
sciences

En bref

- › Langue(s) d'enseignement: Français
- › Ouvert aux étudiants en échange: Oui

Présentation

Description

Dans le contexte de l'apprentissage automatique, l'apprentissage supervisé est souvent privilégié pour sa capacité à prédire des étiquettes ou des valeurs cibles à partir de données d'entraînement étiquetées. Cependant, ces données sont souvent difficiles à obtenir, et les modèles qui en résultent sont limités par la qualité et le domaine de cette supervision. Lorsque l'étiquetage est trop complexe ou trop coûteux à collecter, au regard de la variété des situations auxquelles l'agent décisionnel peut être confronté, l'apprentissage par renforcement offre une alternative prometteuse. L'apprentissage par renforcement, contrairement à l'apprentissage supervisé, n'exige pas une supervision explicite pour chaque exemple de données. Au lieu de cela, l'agent apprend au travers d'interactions avec son environnement, sur la base de récompenses collectées en fonction de l'utilité de ses actions pour la tâche visée.

Cette capacité à apprendre par l'expérience rend l'apprentissage par renforcement applicable à un bien plus large spectre d'applications, et permet de connecter plus finement les objectifs d'apprentissage optimisés avec les objectifs souhaités pour le modèle résultant.

L'apprentissage par renforcement est particulièrement bien adapté aux problèmes de prise de décision séquentielle, où les actions prises à un moment donné affectent les récompenses futures. Par exemple, dans le domaine de la robotique, un robot peut apprendre à naviguer dans un environnement en effectuant des actions et en observant les récompenses associées à ces actions, telles que la distance parcourue ou la minimisation des collisions. C'est également le cas pour des problèmes de génération de données séquentielles, tels que la génération de langue par exemple, où les modèles de générations pre-appris de manière auto-supervisée à prédire le prochain mot en fonction des précédents peuvent être spécialisés pour des tâches particulières par apprentissage par renforcement, permettant de considérer des scores globaux que les séquences produites (e.g. ChatGPT a été finetuné par apprentissage par renforcement sur des préférences d'utilisateurs).

Ce cours vise à couvrir les principaux concepts et méthodes de ce paradigme d'apprentissage en pleine expansion, en mettant l'accent sur les différentes approches et leurs applications pratiques. Il pourra se structurer de la manière suivante :

1. Introduction à l'Apprentissage par Renforcement

- Comparaison avec l'apprentissage supervisé
- Processus de Décision de Markov (MDP)
- Méthodes Tabulaires de Planification
- Algorithme de Bellman
- Policy Iteration et Value Iteration

2. Méthodes basées sur des Valeurs

- Q-Learning, Sarsa
- Deep Q-Networks (DQN)

3. Apprentissage par Policy Gradient et Méthodes Actor-Critic

- Policy Gradient
- Méthodes Actor-Critic: A2C, TRPO, PPO,

4. Apprentissage Hors Politique et Méthodes pour Actions Continues

- Méthodes hors politique
- Méthodes pour actions continues (DDPG, SAC, TD3)

Objectifs

- Comprendre les différences de paradigme entre apprentissage supervisé et apprentissage par renforcement,
- Savoir formaliser des problèmes décisionnels sous la forme de problèmes d'apprentissage par renforcement
- Savoir implémenter les principaux algorithmes (et les faire fonctionner sur des environnements simples GYM)

Heures d'enseignement

CM	Cours magistral	6h
TP	Travaux pratique	8h

Pré-requis obligatoires

Solide bases d'apprentissage statistique, en particulier supervisé

- Notions sur l'apprentissage profond
- Connaissances en programmation Numpy et PyTorch

Infos pratiques

Lieu(x)

> Angers

Campus

› Campus Belle-beille