Conservatoire national des arts et métiers

USEET8 - Reinforcement Learning

Présentation

Prérequis

- Students are required to have taken an introductory machine learning course.
- Good knowledge on probability and statistics is expected.
- Bases on Markov Chains are recommended, but this is not a prerequisite.

Objectifs pédagogiques

This course provides an overview of reinforcement learning (RL) methods. Both theoretical and programming aspects will be extensively explored in this course in order to acquire a solid expertise on both. By the end of the course, students should:

- Understand the notion of stochastic approximations and their relation with RL;
- Understand the basis of Markov decision theory;
- Apply Dynamic Programming methods to solve the Bellman equations;
- Master the basic techniques of Reinforcement Learning: Monte Carlo, Time-difference and Policy Gradient;
- Study a proof of convergence for RL algorithms;
- Master more advanced techniques such as actor-critic methods and deep RL.

Programme

Contenu

This course will introduce machine learning techniques based on stochastic approximations and MDP models, i.e., SARSA, Q-learning, policy gradient. Two homework assignments will focus on implementing these techniques, in order to learn how to master them by direct implementation. A project in teams of 2/3 students will permit to address more advanced techniques and problems in the field of RL and more in general the application of Markov theory for modeling and optimization.

Lectures:

- Course Overview. Introduction to Markov decision theory, stochastic approximations, and reinforcement learning;
- Stochastic approximations: the Robbins-Monro algorithm;
- Criteria for convergence;
- Application to admission control problems;
- Markov decision processes: definitions, average cost and discounted cost;
- Bellman equations. Solutions based on Dynamic Programming;
- Monte Carlo methods for Reinforcement Learning;
- Time Difference methods: SARSA and Q-Learning;
- Proof of convergence of Q-Learning;
- Policy gradient: REINFORCE;
- Actor-critic methods;
- Multi-armed bandits;
- Deep-reinforcement Learning.

Lab assignments:

- Practice of stochastic approximation on a traffics admission problem;
- Practice of Montecarlo, Q-learning and SARSA on gridworld (discounted cost);
- Practice of buffer management with admission control (average cost).

Modalités de validation

- Contrôle continu
- Projet(s)



🌞 Mis à jour le 03-07-2024

Code : USEET8

Unité spécifique de type cours 3 crédits

Responsabilité nationale : EPN05 - Informatique / Stefano SECCI

- Mémoire
- Examen final

Description des modalités de validation

Final exam, lab and research project reports.

All students in the class will also conduct a research project in the field of reinforcement learning and write a short 5-page paper. Subjects will be provided during the first-class session, related to Constrained RL and Delayed RL.

Bibliographie

Titre	Auteur(s)
Artificial Intelligence: A modern approach, 3rd edition, 2010.	S. Russell, P. Norvig, Prentice Hall
Reinforcement Learning: An Introduction, 1992	R. S. Sutton, A. G. Barto, MIT Press