

Guia docent

300119 - REINFL - Aprenentatge per Reforçament

Última modificació: 04/07/2026

Unitat responsable: Escola d'Enginyeria de Telecomunicació i Aeroespacial de Castelldefels
Unitat que imparteix: **Titulació:** MÀSTER UNIVERSITARI EN INTEL·LIGÈNCIA ARTIFICIAL PER A INDÚSTRIES
CONNECTADES (AI4CI) (Pla 2025). (Assignatura optativa).
Curs: 2026 **Crèdits ECTS:** 3.0 **Idiomes:** Anglès

PROFESSORAT

Professorat responsable: Francesco de Pellegrini (Univ. Avignon)

Altres: Francesco de Pellegrini (Univ. Avignon)

CAPACITATS PRÈVIES

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 (see the course "Performance Evaluation of Connected Systems") but this is not a prerequisite.

REQUISITS

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 (see the course "Performance Evaluation of Connected Systems") but this is not a prerequisite.

METODOLOGIES DOCENTS

Lectures, lab, assignments, project

OBJECTIUS D'APRENTATGE DE L'ASSIGNATURA

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.

HORES TOTS DE DEDICACIÓ DE L'ESTUDIANTAT

Tipus	Hores	Percentatge
Hores grup gran	45,0	36.00
Hores aprenentatge autònom	80,0	64.00

Dedicació total: 125 h

CONTINGUTS

Reinforcement Learning

Descripció:

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.

1- 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.

2- 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).

Dedicació: 75h

Grup gran/Teoria: 27h

Aprenentatge autònom: 48h



SISTEMA DE QUALIFICACIÓ

Qualsevol acte de frau acadèmic, plagi, ús o tinença a l'abast de mitjans no autoritzats en qualsevol activitat d'avaluació comportarà la qualificació de zero (0) en la prova o lliurament afectat. A més, d'acord amb la normativa de la Universitat, la possible derivació dels fets per a l'obertura d'un expedient disciplinari implicarà que l'assignatura quedi en l'estat provisional de "pendent d'avaluació" fins a la resolució de l'expedient. La gestió d'aquestes incidències es desplega d'acord amb el [Marc d'actuació per a la integritat acadèmica en l'avaluació de la UPC](#).

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.