

Guía docente

300119 - REINFL - Aprendizaje por Refuerzo

Última modificación: 04/07/2026

Unidad responsable: Escuela de Ingeniería de Telecomunicación y Aeroespacial de Castelldefels
Unidad que imparte: **Titulación:** MÁSTER UNIVERSITARIO EN INTELIGENCIA ARTIFICIAL PARA INDUSTRIAS CONECTADAS (AI4CI) (Plan 2025). (Asignatura optativa).

Curso: 2026 **Créditos ECTS:** 3.0 **Idiomas:** Inglés

PROFESORADO

Profesorado responsable: Francesco de Pellegrini (Univ. Avignon)

Otros: Francesco de Pellegrini (Univ. Avignon)

CAPACIDADES PREVIAS

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.

REQUISITOS

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.

METODOLOGÍAS DOCENTES

Lectures, lab, assignments, project

OBJETIVOS DE APRENDIZAJE DE LA ASIGNATURA

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.

HORAS TOTALES DE DEDICACIÓN DEL ESTUDIANTE

Tipo	Horas	Porcentaje
Horas aprendizaje autónomo	80,0	64.00
Horas grupo grande	45,0	36.00

Dedicación total: 125 h

CONTENIDOS

Reinforcement Learning

Descripción:

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).

Dedicación: 75h

Grupo grande/Teoría: 27h

Aprendizaje autónomo: 48h

SISTEMA DE CALIFICACIÓN

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.