

---

Curso Académico: ( 2023 / 2024 )

Fecha de revisión: 08-05-2023

---

Departamento asignado a la asignatura: Departamento de Informática

Coordinador/a: FERNANDEZ REBOLLO, FERNANDO

Tipo: Optativa Créditos ECTS : 3.0

Curso : 1 Cuatrimestre : 1

---

#### REQUISITOS (ASIGNATURAS O MATERIAS CUYO CONOCIMIENTO SE PRESUPONE)

Se recomiendan los conocimientos de la asignatura de Aprendizaje Automático

#### DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

Introducción al Aprendizaje por Refuerzo

- Introducción al aprendizaje por refuerzo
- Procesos de Decisión de Markov
- Políticas y optimalidad: horizonte infinito descontado
- Funciones de Valor

Programación Dinámica

- Resolución de problemas sobre MDP: métodos libres de modelo, basados en modelo y programación dinámica
- Algoritmo de Iteración de Política
- Algoritmo de Iteración de Valor

Aprendizaje por refuerzo directo

- Métodos Monte Carlo: y Monte Carlo con arranque exploratorio
- Métodos libre de modelo: Q-Learning
- Ejemplo de ejecución de Q-Learning
- Métodos on-policy vs. off-policy: SARSA
- Exploración y explotación: e-greedy y softmax

Métodos Basados en el Modelo

- Aprendizaje de Modelos
- Dyna-Q

Representación en Aprendizaje por Refuerzo

- Representación del espacio de estados, acciones y Q
- Discretización del espacio de estados: métodos uniformes y adaptativos
- Métodos aproximados para representar la función Q: Batch Q-Learning

Generalización Mediante Aproximación de Funciones

- Aproximación mediante redes de neuronas
- Deep reinforcement learning

Métodos de Búsqueda de la Política

- Aproximación de la Política
- Métodos actor-crítico
- Proximal Policy Optimization (PPO)

Otros temas de Aprendizaje por Refuerzo

- Aprendizaje por Refuerzo jerárquico
- Transferencia de aprendizaje aprendido
- Aprendizaje por Refuerzo multi-agente
- Aprendizaje por Refuerzo seguro
- Aprendizaje por Refuerzo fuera de línea
- Aprendizaje por Refuerzo multi-objetivo

- Aprendizaje por Refuerzo parcialmente observable

Aprendizaje por Refuerzo en el mundo real:

- Aplicaciones del aprendizaje por refuerzo
- Marcos y software de aprendizaje por refuerzo

## SISTEMA DE EVALUACIÓN

La evaluación de esta asignatura se basa en tres elementos principales:

- Examen final: 30% de la calificación final. El examen se compondrá de una serie de preguntas teóricas de respuesta corta y/o tipo test
- Trabajos del curso: 60% de la calificación final. Desarrollo de un proyecto práctico a lo largo del curso a realizar en grupo
- Participación en clase: 10 % de la calificación final. Se centra principalmente en la exposición de un tema de la asignatura

La calificación final se obtendrá como una suma ponderada de las calificaciones de cada apartado. No hay nota mínima en ningún apartado.

**Peso porcentual del Examen Final:** 30

**Peso porcentual del resto de la evaluación:** 70

## BIBLIOGRAFÍA BÁSICA

- Richard Sutton and Andrew Barto Reinforcement Learning: an Introduction, The MIT Press.

## RECURSOS ELECTRÓNICOS BÁSICOS

- DeepMind . MuJoCo: <https://mujoco.org/>
- Open AI . Open AI Proximal Policy Optimization: <https://openai.com/research/openai-baselines-ppo>
- Open AI . Gymnasium: <https://gymnasium.farama.org/>