



FICHA TÉCNICA DE LA ASIGNATURA

Datos de la asignatura	
Nombre completo	Aprendizaje por Refuerzo
Código	DEAC-IMAT-411
Título	Grado en Ingeniería Matemática e Inteligencia Artificial
Impartido en	Grado en Ingeniería Matemática e Inteligencia Artificial [Cuarto Curso]
Nivel	Reglada Grado Europeo
Cuatrimestre	Semestral
Créditos	3,0 ECTS
Carácter	Obligatoria (Grado)
Departamento / Área	Departamento de Electrónica, Automática y Comunicaciones
Responsable	Álvaro Jesús López López
Horario de tutorías	Concertar cita por correo electrónico

Datos del profesorado	
Profesor	
Nombre	Álvaro Jesús López López
Departamento / Área	Instituto de Investigación Tecnológica (IIT)
Despacho	Francisco de Ricci, 3
Correo electrónico	Alvaro.Lopez@iit.comillas.edu
Teléfono	4503
Profesores de laboratorio	
Profesor	
Nombre	Lucía Güitta López
Departamento / Área	Instituto de Investigación Tecnológica (IIT)
Despacho	Francisco de Ricci, 3
Correo electrónico	lguita@comillas.edu
Teléfono	4518
Profesor	
Nombre	Rodrigo Sánchez Molina
Departamento / Área	Departamento de Electrónica, Automática y Comunicaciones
Correo electrónico	rsmolina@icai.comillas.edu

DATOS ESPECÍFICOS DE LA ASIGNATURA

Contextualización de la asignatura



Aportación al perfil profesional de la titulación

El aprendizaje por refuerzo (RL, por sus siglas en inglés) es una forma de aprendizaje automático basada en el principio de prueba y error. A diferencia de los métodos tradicionales de aprendizaje supervisado y no supervisado, donde los algoritmos se basan en datos etiquetados o en la identificación de patrones subyacentes en los datos, el aprendizaje por refuerzo se centra en cómo los agentes pueden aprender comportamientos óptimos a través de la interacción directa con su entorno. Esta técnica ha demostrado ser altamente efectiva tanto de forma aislada como en combinación con otras técnicas de aprendizaje automático, ofreciendo una versatilidad que permite su aplicación en diversos campos como la robótica, el control de sistemas, los videojuegos, y la optimización de decisiones en tiempo real.

Al completar esta asignatura, los alumnos estarán familiarizados con conceptos fundamentales del aprendizaje por refuerzo, tales como la recompensa, el retorno, las funciones de valor, las políticas, los estados y la aproximación de estados. Este conocimiento permitirá a los estudiantes comprender cómo diseñar y desarrollar agentes artificiales que puedan exhibir comportamientos inteligentes en entornos no estacionarios (cambiantes). Además, el curso ofrece un equilibrio entre la teoría y la práctica, proporcionando a los estudiantes la oportunidad de aplicar estos conceptos en problemas del mundo real, preparando así a los futuros profesionales para enfrentar desafíos complejos en entornos dinámicos y en constante evolución.

Prerrequisitos

Fundamentos de programación en Python

Competencias - Objetivos

Competencias

GENERALES

CG06	Capacidad para utilizar el aprendizaje de manera estratégica y flexible en función del objetivo perseguido, a partir del reconocimiento del propio sistema de aprendizaje y de la conciencia del aprendizaje mismo, dentro de un contexto tecnológico que evoluciona rápidamente
CG09	Capacidad para determinar eficazmente los objetivos, prioridades, métodos y controles para desempeñar tareas relacionadas con la planificación de proyectos de explotación de datos e inteligencia artificial, mediante la organización de las actividades con los plazos y los medios disponibles
CG14	Capacidad para integrar conocimiento multidisciplinar en un determinado proyecto o sistema.

ESPECÍFICAS

CE01	Capacidad para la resolución de los problemas matemáticos que puedan plantearse en la ingeniería, aplicando con aptitud los conocimientos sobre: álgebra lineal y multilineal, geometría, cálculo diferencial e integral, ecuaciones diferenciales, métodos numéricos, estadística y optimización.
CE04	Capacidad para utilizar con habilidad y soltura software matemático, así como para implementar algoritmos y desarrollar programas informáticos que permitan resolver los problemas matemáticos planteados en el ámbito de la ingeniería y de la inteligencia artificial.
CE07	Aptitud para modelar y resolver sistemas físicos en el ámbito de la ingeniería, mediante técnicas de cálculo numérico, álgebra numérica, ecuaciones en diferencias, ecuaciones diferenciales o técnicas propias de la matemática discreta.



CE10	Conocimiento de la sintaxis, las estructuras principales y los elementos básicos de un lenguaje de programación en el contexto del análisis de datos y la inteligencia artificial
CE25	Conocimiento y capacidad para aplicar técnicas de inteligencia artificial, aprendizaje automático, aprendizaje profundo y aprendizaje por refuerzo que permiten extraer conocimiento de grandes volúmenes de datos.
CE26	Capacidad para aplicar técnicas de inteligencia artificial adecuadas para la realización de trabajos y proyectos de ingeniería.
CE31	Capacidad para especificar, diseñar e implementar las técnicas de aprendizaje automático y profundo para la resolución de problemas complejos.
CE33	Capacidad para analizar el comportamiento de los sistemas físicos en el dominio del tiempo. Conocimiento de los principios de los sistemas de control en lazo cerrado: estabilidad, precisión, rapidez y amortiguamiento.

Resultados de Aprendizaje

RA1	Comprender las técnicas que permiten a los agentes artificiales adaptarse a entornos cambiantes
RA2	Comprender la importancia de explorar adecuadamente el espacio de operación, así como las fórmulas habituales para conseguirlo
RA3	Conocer y comprender la estructura de un proceso de decisión Markoviano, con especial énfasis en las fórmulas recursivas que permiten asignar valor a estados y acciones
RA4	Distinguir entre los tres grandes problemas que se definen en el marco de los MDPs con el objetivo de operar de manera óptima: predicción de las funciones de valor, control o maximización de dichas funciones, y planificación (entendida como la solución del problema de control usando un modelo del entorno)
RA5	Conocer las técnicas de programación dinámica para solución de los problemas de predicción y control, así como de los importantes teoremas en los que se apoyan
RA6	Conocer las técnicas de Monte Carlo para la solución aproximada de los problemas de predicción y control, así como la estructura y funcionamiento dinámico de los algoritmos de Monte Carlo RL
RA7	Conocer las técnicas de diferencia temporal (TD learning) para la solución aproximada de los problemas de predicción y control, así como la estructura y funcionamiento dinámico de los principales algoritmos de TD learning (SARSA y Q-learning).
RA8	Entender las limitaciones del enfoque tabular, así como las técnicas de aproximación empleadas para generalizar los métodos de RL
RA9	Comprender las implicaciones de la generalización y la discriminación entre estados de los agentes artificiales
RA10	Conocer las distintas aproximaciones para encontrar una política optimizada mediante RL, comprendiendo las ventajas y desventajas de cada una en términos de sesgo y varianza durante el aprendizaje

BLOQUES TEMÁTICOS Y CONTENIDOS



Contenidos – Bloques Temáticos

TEORÍA

Tema 1: Introducción

- Evolución histórica del aprendizaje por refuerzo
- Conceptos clave: recompensa, valor de acciones, equilibrio entre exploración y explotación

Tema 2: Procesos de decisión markovianos

- Retorno
- Funciones de valor de estado y valor de acción
- Ecuaciones de Bellman

Tema 3: Programación dinámica

- Definición de problemas de predicción, control y planificación
- Ecuaciones óptimas de Bellman
- Teorema de mejora de la política
- Algoritmo de iteración de política
- Algoritmo de iteración de valor
- Algoritmos de mejora general de política (Generalized Policy Improvement)

Tema 4: Aprendizaje por refuerzo sin modelo

- Aproximación de Monte Carlo
- Aproximación de diferencia temporal (TD learning)
- Fórmulas intermedias
- Aprendizaje on-policy y off-policy
- Descripción de los principales algoritmos de aprendizaje TD basados en funciones de valor (SARSA, Q-learning)

Tema 5: Representación aproximada del entorno

- Tile coding
- Métodos lineales
- Métodos no lineales
- Aproximación directa de la política
- Métodos de actor y crítico

Tema 6: Tendencias actuales

- Descripción de los principales algoritmos de aprendizaje que componen el estado actual de la técnica: DQN, DPG, PPO, A2C, TD3, RLHF

LABORATORIO

Práctica 1

Optimización de una trayectoria en un problema tabular mediante técnicas de programación dinámica



Práctica 2

Optimización de una trayectoria en un problema tabular mediante técnicas de aprendizaje por refuerzo sin modelo

Práctica 3 (proyecto)

Diseño de un agente artificial para optimizar una operación logística en un almacén

METODOLOGÍA DOCENTE

Aspectos metodológicos generales de la asignatura

Metodología Presencial: Actividades

Clase magistral y presentaciones generales. El profesor explicará los conceptos fundamentales de cada tema incidiendo en lo más importante y a continuación se explicarán una serie de problemas tipo, gracias a los cuáles se aprenderá a identificar los elementos esenciales del planteamiento y la resolución de problemas del tema.

Clase invertida. En estas sesiones se trabajará en el refuerzo de conceptos cuyos materiales han sido previamente compartidos por los alumnos y estudiados por ellos.

Resolución de problemas de carácter práctico o aplicado. En estas sesiones se explicarán, corregirán y analizarán problemas de cada tema previamente propuestos por el profesor y trabajados por el alumno.

Prácticas de laboratorio. Se realizarán en grupos y en ellas los alumnos ejercitarán los conceptos y técnicas estudiadas, familiarizándose con el entorno material y humano del trabajo en el laboratorio.

Metodología No presencial: Actividades

Estudio personal. El alumno debe realizar un trabajo personal posterior a las clases teóricas para comprender e interiorizar los conocimientos aportados en la materia.

Ejercicios prácticos y resolución de problemas. Una vez estudiados los conceptos teóricos, el alumno debe ponerlos en práctica para resolver problemas. Pasado un cierto tiempo desde su planteamiento, dispondrá de la solución y podrá solicitar tutorías con el profesor si lo necesita.

Sesiones prácticas de laboratorio. Las prácticas de laboratorio requerirán la realización de un trabajo previo de preparación y finalizarán con la redacción de un informe.

Proyecto. Se presentará un problema general de aprendizaje por refuerzo a todos los alumnos. Los alumnos deberán diseñar en grupo su agente aplicando las técnicas aprendidas con el objetivo de optimizar una determinada operación. El proyecto comprenderá tanto trabajo en las sesiones presenciales como trabajo en el tiempo entre las mismas.



RESUMEN HORAS DE TRABAJO DEL ALUMNO

HORAS PRESENCIALES				
Clases magistrales expositivas y participativas	Sesiones prácticas con uso de software	Tutorías para resolución de dudas	Actividades de evaluación continua del rendimiento	Presentaciones orales
13.00	8.00	5.00	1.00	2.00
HORAS NO PRESENCIALES				
Sesiones prácticas con uso de software	Estudio personal	Proyectos		
8.00	21.00	30.00		
CRÉDITOS ECTS: 3,0 (88,00 horas)				

EVALUACIÓN Y CRITERIOS DE CALIFICACIÓN

Actividades de evaluación	Criterios de evaluación	Peso
<ul style="list-style-type: none">• Prueba intersemestral• Examen final	<ul style="list-style-type: none">• Comprensión de los conceptos teóricos• Aplicación de estos conceptos a la resolución de problemas prácticos• Análisis crítico de los resultados numéricos• Comunicación escrita	50 %
<ul style="list-style-type: none">• Preparación de prácticas de laboratorio• Redacción de informes• Participación activa	<ul style="list-style-type: none">• Comprensión de conceptos.• Aplicación de conceptos a la resolución de problemas prácticos y a la realización de prácticas en el laboratorio.• Análisis e interpretación de los resultados obtenidos en las prácticas de laboratorio.• Capacidad de trabajo en grupo.• Presentación y comunicación escrita.	10 %
<ul style="list-style-type: none">• Proyecto	<ul style="list-style-type: none">• Calidad del diseño• Sentido crítico en la aplicación de conceptos teóricos al diseño del agente• Uso de funciones avanzadas de aprendizaje por refuerzo• Monitorización del proceso de aprendizaje y del funcionamiento en operación del agente• Autonomía y habilidad para resolver problemas• Comunicación escrita	30 %
	<ul style="list-style-type: none">• Capacidad de síntesis	



- Presentación del proyecto

- Sentido crítico para poner el foco en lo importante
- Comunicación oral

10 %

Calificaciones

En la convocatoria extraordinaria se realizan un nuevo examen final y otro de laboratorio, este último solo si estuviera suspenso. Se conservan las notas de la convocatoria ordinaria de todas aquellas actividades de evaluación que no deban repetirse. La calificación final se obtendrá de la misma forma que en la convocatoria ordinaria, pero sustituyendo la nota de laboratorio por la del examen extraordinario en caso de que este bloque estuviera suspenso en la convocatoria ordinaria.

CONVOCATORIA ORDINARIA

El peso de cada una de las actividades de evaluación será el siguiente:

- Teoría (50 %)
 - Prueba intersemestral: 15 %
 - Examen final: 35 %
- Laboratorio (50 %)
 - Prácticas: 10 %
 - Informe proyecto: 30 %
 - Presentación proyecto: 10 %

La calificación final se calculará atendiendo a las siguientes restricciones:

- La nota de teoría será la media ponderada de la prueba intersemestral y el examen final, siempre y cuando la nota del examen final sea mayor que 4. En caso contrario, la nota de teoría será la mínima entre la media ponderada de ambas pruebas y la del examen final.
- La nota de laboratorio será la media ponderada de las tres calificaciones, siempre y cuando la nota del informe del proyecto sea mayor o igual que 5.

CONVOCATORIA EXTRAORDINARIA

En la convocatoria extraordinaria se realizan un nuevo examen final y otro de laboratorio, este último solo si estuviera suspenso.

Se conservan las notas de la convocatoria ordinaria de todas aquellas actividades de evaluación que no deban repetirse.

La calificación final se obtendrá de la misma forma que en la convocatoria ordinaria, atendiendo a las mismas restricciones en la parte de teoría pero sustituyendo la nota de laboratorio por la del examen de la convocatoria extraordinaria en caso de que haya tenido que realizarse.

PLAN DE TRABAJO Y CRONOGRAMA

Actividades	Fecha de realización	Fecha de entrega
Sesiones de teoría	Semanas 1, 2, 4, 6, 8, 10, 12	



COMILLAS

UNIVERSIDAD PONTIFICIA

ICAI

ICADE

CIHS

GUÍA DOCENTE
2024 - 2025

Prácticas de laboratorio	Semanas 3 y 5	
Prácticas de proyecto	Semanas 7 y 9	
Presentación oral proyecto	Semana 11	

BIBLIOGRAFÍA Y RECURSOS

Bibliografía Básica

"Reinforcement Learning: An Introduction". R. Sutton, and A. Barto. The MIT Press, Second edition, (2018)

En cumplimiento de la normativa vigente en materia de **protección de datos de carácter personal**, le informamos y recordamos que puede consultar los aspectos relativos a privacidad y protección de datos que ha aceptado en su matrícula entrando en esta web y pulsando "descargar"

[https://servicios.upcomillas.es/sedelectronica/inicio.aspx?csv=02E4557CAA66F4A81663AD10CED66792](https://servicios.upcomillas.es/sedeelectronica/inicio.aspx?csv=02E4557CAA66F4A81663AD10CED66792)