

Formulario de aprobación de curso de posgrado/educación permanente

Asignatura: Taller de Aprendizaje por Refuerzo

(Si el nombre contiene siglas deberán ser aclaradas)

Modalidad:

(posgrado, educación permanente o ambas)

Posgrado

Educación permanente

Profesor de la asignatura ¹: Isabel Amigo (gr3, IIE)

(título, nombre, grado o cargo, instituto o institución)

Profesor Responsable Local ¹:

(título, nombre, grado, instituto)

Otros docentes de la Facultad: Federico La Rocca (gr4,IIE), Mariana del Castillo (gr2, IIE)

(título, nombre, grado, instituto)

Docentes fuera de Facultad:

(título, nombre, cargo, institución, país)

¹ Agregar CV si el curso se dicta por primera vez.

(Si el profesor de la asignatura no es docente de la Facultad se deberá designar un responsable local)

Programa(s) de posgrado: Maestría en Ingeniería Matemática (IngeMat), Maestría en Ciencia de Datos (MCDAA) y Aprendizaje Automático, Maestría y Doctorado en Ingeniería Eléctrica

Instituto o unidad: IIE

Departamento o área:

Horas Presenciales: 40

(se deberán discriminar las horas en el ítem Metodología de enseñanza)

Nº de Créditos: 8

[Exclusivamente para curso de posgrado]

(de acuerdo a la definición de la UdelaR, un crédito equivale a 15 horas de dedicación del estudiante según se detalla en el ítem Metodología de enseñanza)

Público objetivo: Estudiantes de posgrado (y eventualmente estudiantes avanzados de grado) con formación básica en álgebra lineal y probabilidad y cuyos intereses estén relacionados con el

modelado y control de sistemas complejos desde una perspectiva de aprendizaje automático. El curso espera llegar a estudiantes de programas de la Facultad de Ingeniería (IE, IngeMat, MCDAA) pero también a estudiantes de posgrado por fuera de la facultad como la maestría en Informática del PEDECIBA.

Cupos: No tiene cupos

Objetivos: Aprender qué son los procesos de decisión markoviana (MDP) y cómo se utilizan para modelar problemas de aprendizaje por refuerzo. Aprender los métodos básicos de resolución de MDPs y algoritmos de resolución de problemas de Aprendizaje por Refuerzo. Familiarizarse con la programación de dichos algoritmos y con la utilización de librerías de software modernas para su resolución.

Conocimientos previos exigidos: Álgebra de matrices. Fundamentos de probabilidad y estadística. Programación a nivel básico.

Conocimientos previos recomendados: Programación en python científico.

Metodología de enseñanza:

(comprende una descripción de la metodología de enseñanza y de las horas dedicadas por el estudiante a la asignatura, distribuidas en horas presenciales -de clase práctica, teórico, laboratorio, consulta, etc.- y no presenciales de trabajo personal del estudiante)

Descripción de la metodología:

El curso consta de una clase teórica presencial durante 10 semanas además de 5 laboratorios asociados a los temas del curso.

Los laboratorios incluyen un taller práctico en python que busca ilustrar algunos de los conceptos vistos en clase. Éstos serán realizados por los estudiantes y servirán sobre todo para auto-estudio y preparación del trabajo final. El objetivo es que luego de realizar el laboratorio el estudiante sea capaz de modelar y programar problemas bajo la forma de un MDP y como un problema de RL, y resolver los mismos a través de la programación de algoritmos y de la utilización de bibliotecas.

Detalle de horas:

- Horas de clase (teórico): 30
- Horas de consulta: 10

- Horas de evaluación:
 - Subtotal de horas presenciales: 40
 - Horas de estudio: 15
 - Horas de resolución de laboratorios: 20
 - Horas proyecto final/monografía: 50
 - Total de horas de dedicación del estudiante: 125
-

Forma de evaluación:

Entrega de ejercicios prácticos y/o en el lenguaje de programación python junto con un proyecto final.

En el proyecto final, los estudiantes podrán investigar y aplicar algoritmos de aprendizaje por refuerzo del arte, a una aplicación de su interés o replicar resultados de artículos científicos.

Una vez concluido el componente instructivo del curso y realizada la entrega de ejercicios, los estudiantes deberán seleccionar un tema específico, relacionado con el material tratado en clase, y realizar un estudio relativamente profundo del tema. Esto implica por ejemplo buscar fuentes de literatura (trabajos de investigación, libros de texto, etc.), resolver el problema programando y realizar una demostración y analizar los resultados. La entrega final consistirá en un resumen detallado de las ideas principales en formato paper.

Temario:

1. Introducción y motivación
 2. MDPs
 - a) Función de valor
 - b) Ecuaciones de Bellman
 - c) Algoritmo de iteración de valor (value iteration)
 - d) Algoritmo de iteración en la política (policy iteration)
 3. Aprendizaje por refuerzo
 4. Algoritmos de resolución
 - a) TD(0)
 - b) SARSA
 - c) Q-learning
 5. Aproximación de función
-

6. Gradiente estocástico
 7. Aprendizaje profundo (Deep RL)
-

Bibliografía:

Reinforcement Learning: An Introduction, Richard S. Sutton and Andrew G. Barto
Second Edition, MIT Press, Cambridge, MA, 2018
Reinforcement Learning and Stochastic Optimization: A Unified Framework for Sequential Decisions, Warren B. Powell, Wiley, 2022
Algorithms for Reinforcement Learning, Csaba Szepesvári, Morgan and Claypool Publishers, 2010.
Reinforcement Learning and Optimal Control, Dimitri Bertsekas, Athena Scientific, 2019
Grokking Deep Reinforcement Learning, Miguel Morales, Manning, 2020.
Deep Reinforcement Learning Hands-On (2a edición), Maxim Lapan, Packt Publishing, 2020.

Datos del curso

Fecha de inicio y finalización: 14/04/2025 - 4/07/2025

Horario y Salón: a determinar con los estudiantes

Arancel:

[Si la modalidad no corresponde indique "no corresponde". Si el curso contempla otorgar becas, indíquelo]

Arancel para estudiantes inscriptos en la modalidad posgrado: no corresponde

Arancel para estudiantes inscriptos en la modalidad educación permanente:

4000 UI

El curso contempla otorgar becas a solicitud de los estudiantes.
