



1821 Universidad de Buenos Aires

Resolución Consejo Directivo

Número:

Referencia: EX-2025-03588385- -UBA-DMESA#FCEN - POSTGRADO - Sesión
18/08/2025

VISTO:

La nota presentada por la Dirección del Departamento de Matemática, mediante la cual eleva la información del curso de posgrado **Aprendizaje por Refuerzo y Garantías Teóricas** para el año 2025,

CONSIDERANDO:

lo actuado por la Comisión de Doctorado,

lo actuado por este Cuerpo en la sesión realizada el día 18 de agosto de 2025,

en uso de las atribuciones que le confiere el Artículo 113° del Estatuto Universitario,

**EL CONSEJO DIRECTIVO DE LA FACULTAD
DE CIENCIAS EXACTAS Y NATURALES**

RESUELVE:

ARTÍCULO 1°: Aprobar el nuevo curso de posgrado **Aprendizaje por Refuerzo y Garantías Teóricas** de 24 horas y 4 semanas de duración, que será dictado por la Dra. Gabriela Jerónimo, con la colaboración del Prof. Matthieu Jonckheere.

ARTÍCULO 2°: Aprobar el programa del curso de posgrado **Aprendizaje por Refuerzo y Garantías Teóricas** que como anexo forma parte de la presente Resolución, para su dictado en octubre de 2025.

ARTÍCULO 3°: Aprobar un puntaje máximo de un (1) punto para la Carrera de Doctorado.

ARTÍCULO 4°: Establecer un arancel de **CATEGORÍA NULA**.

ARTÍCULO 5°: Disponer que, de no mediar modificaciones en el programa, la carga horaria y el arancel, el presente Curso de Posgrado tendrá una vigencia de cinco (5) años a partir de la fecha de la presente Resolución.

ARTÍCULO 6°: Comuníquese a todos los Departamentos Docentes, a la Dirección de Estudiantes y Graduados, a la Biblioteca de la FCEyN y a la Secretaría de Posgrado con copia del programa incluida. Cumplido, pase a MATEMATICA#FCEN y resérvese.

ANEXO

Aprendizaje por Refuerzo y Garantías Teóricas

PROGRAMA

Objetivos:

El curso tiene como objetivo introducir a los estudiantes en los conceptos subyacentes de los algoritmos de aprendizaje por refuerzo. Más allá de la teoría, los estudiantes tendrán la oportunidad de practicar el Aprendizaje por Refuerzo a través de sesiones de laboratorio.

Contenidos:

Los métodos de Aprendizaje por Refuerzo (RL) se encuentran en la intersección de varios campos como el aprendizaje automático, la teoría de probabilidades, la teoría de control y la teoría de juegos. Estudian los compromisos fundamentales entre la explotación y la exploración para optimizar una recompensa futura en un entorno desconocido. Recientemente, han sido aplicados con éxito en muchas aplicaciones como juegos complejos (Go o Atari), control de robots, selección de expertos, logrando en muchos casos un rendimiento sobrehumano. Esta clase cubrirá las herramientas necesarias (desde un punto de vista teórico y práctico) para entender estos avances.

Programa:

1. Introducción al Aprendizaje por Refuerzo (RL) y breve introducción a cadenas de Markov.

- Diferencias con otras áreas de ML
- Análisis markoviano en espacio de estados finito
- Análisis del valor medio
- Ley de los grandes números (LLN), convergencia al régimen estacionario

2. Procesos de decisión Markovianos

- Ecuaciones de Bellman
- Programación dinámica
- TP1 sobre estas nociones

3. Problemas de Bandits.

- Primeros ejemplos de compromiso entre explotación y exploración
- Algoritmos clásicos
- TP2 sobre Bandits

4. Problemas tabulares sin modelo.

- Monte Carlo
- TD(λ)
- SARSA, Q-learning
- Elementos de aproximación estocástica, demostración de propiedades de convergencia
TP3 sobre Q-learning

5 .Espacios de estado más grandes.

- Aproximaciones de funciones de valor
- Aproximaciones de políticas
- Gradiente de política, Lema de evaluación
- Estimación del gradiente, exploración con recompensas raras
- Proyecto

BIBLIOGRAFIA

- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction (2nd ed.)*. MIT Press.
- Bertsekas, D. P. (2017). *Dynamic Programming and Optimal Control (4th ed.)*. Athena Scientific, Vols I & II.
- Puterman, M. L. (2014). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons.
- *Score-Aware Policy-Gradient Methods and Performance Guarantees using Local Lyapunov Conditions: Applications to Product-Form Stochastic Networks and Queueing Systems* C Comte, M. Jonckheere J Sanders, A Senen-Cerda, *JMLR 2025 (to be published)*
- *Symphony of experts: orchestration with adversarial insights in reinforcement learni* M. Jonckheere, C Mignacco, G Stoltz, *TMLR 2025*

