

Aprendizaje por Refuerzo

Nombre de la asignatura:	Aprendizaje por refuerzo			
Tipo de asignatura:	Asignatura optativa			
Tiempo de dedicación del estudiante a las actividades de:				
DOC	TIS	TPS	Horas totales	Créditos
48	20	100	168	6

1. Historial de la asignatura.

Fecha revisión/actualización	Participantes	Observaciones, cambios y/o justificación
Enero de 2026	Dr. Julio Alberto García Rodríguez Consejo de Posgrado del Departamento de Sistemas y Computación	

2. Prerrequisitos y correquisitos.

Probabilidad y Estadística.

3. Objetivo de la asignatura.

Conocer y aplicar la metodología de aprendizaje de un agente por medio de Aprendizaje por refuerzo, reconociendo el tipo de problema y eligiendo la estrategia más adecuada para resolverlo.

4. Aportación al perfil del graduado.

Esta asignatura aporta al perfil del egresado la habilidad para diseñar e implementar sistemas computacionales inteligentes que aprenden a tomar decisiones óptimas mediante la interacción con su entorno. Además, capacita al estudiante para modelar problemas bajo incertidumbre, generando propuestas aplicables a la investigación y desarrollo tecnológico.

5. Contenido temático.

Unidad	Temas	Subtemas
1	Introducción	1.1 Introducción al Aprendizaje por refuerzo. 1.2 El rol de la probabilidad y la estadística. 1.3 Interacción Agente – Entorno. 1.4 Un tercer paradigma del Aprendizaje de Máquinas.
2	Exploración y Explotación	2.1 El problema del bandido multiarmado. 2.2 Métodos de Acción-Valor. 2.3 Estrategias epsilon-greedy. 2.4 Valores iniciales optimistas. 2.5 Selección de acciones con límites UCB.
3	Markov Decision Process (MDP)	3.1 Interfaz Agente – Ambiente. 3.2 La propiedad de Markov. 3.3 Objetivos y recompensas. 3.4 Regresos y episodios.

4	Programación dinámica	4.1 Evaluación de la política. 4.2 Mejoramiento de la política. 4.3 Iteración de la política. 4.4 Valor de la iteración.
5	Aprendizaje desde la experiencia y escalado	5.1 Predicción en caso libre de modelo. 5.2 Control en caso libre de modelo. 5.3 Integrando aprendizaje y planeación. 5.4 Aproximaciones a la Función Valor. 5.5 Gradientes de la política.

6. Metodología de desarrollo del curso.

- Exposición interactiva apoyada en programas computacionales y simulaciones visuales.
- Integración equipos de trabajo.
- Análisis del Estado del Arte para seleccionar y analizar casos de estudio por tema.
- Promover grupos de discusión para el análisis sobre los conceptos investigados.
- Talleres de implementación de algoritmos en entornos simulados (Gymnasium).
- Aplicación de exámenes.

7. Sugerencias de evaluación.

- Examen escrito.
- Ejercicios resueltos por el alumno.
- Exposición de temas.
- Solución de problemas y casos de estudio.

Exámenes	Tareas / Estudio de Casos / Exposiciones	Participación
50%	30%	20%

8. Prácticas propuestas.

- Ejercicios del libro de texto.
- Implementación de diversos algoritmos relacionados al temario.

9. Libro de texto del curso

- Sutton, R. S., & Barto, A. G. (2018). Reinforcement Learning: An Introduction. Second Edition. MIT Press.
- Lapan, M. (2024). Deep Reinforcement Learning Hands-On: A practical and easy-to-follow guide to RL from Q-learning and DQNs to PPO and RLHF. Third Edition. Packt.
- Morales, M. (2021). Grokking deep reinforcement learning. Manning Publications.

10. Nombre y firma del catedrático responsable.

