



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA
DEPARTAMENTO DE CIENCIA DE LA COMPUTACIÓN

IIC3675 — Aprendizaje Reforzado

Programa de curso

Actualización: 5 de marzo de 2025

Profesor	Rodrigo Toro Icarte (rntoro@uc.cl)
Horario de cátedras	Martes y Jueves, módulo 4
Horario de ayudantías	Viernes, módulo 4
Requisitos	Inteligencia Artificial (IIC2613) o Minería de Datos (IIC2433)

1. Descripción

El aprendizaje reforzado es un área de la inteligencia artificial que estudia la creación de agentes inteligentes. Estos agentes aprenden a resolver problemas mediante ensayo y error. Por ejemplo, es posible dejar a estos agentes jugando videojuegos y, luego de un par de días, aprenden a jugar mejor que los humanos. Este curso enseña las bases teóricas y prácticas de aprendizaje reforzado. Al finalizar el curso, los estudiantes podrán identificar las fortalezas y debilidades de aprendizaje reforzado. También podrán aplicar aprendizaje reforzado a nuevos problemas y proponer nuevos tipos de agentes.

2. Resultados del Aprendizaje

Al finalizar el curso el alumno será capaz de:

1. Evaluar las ventajas y desventajas de utilizar aprendizaje reforzado en un problema dado.
2. Aplicar aprendizaje reforzado para resolver nuevos problemas de decisión exitosamente.
3. Proponer mejoras a agentes existentes de aprendizaje reforzado.
4. Analizar los resultados obtenidos por aprendizaje reforzado para mejorar su rendimiento.

3. Contenidos

1. Introducción al aprendizaje reforzado

- Motivación
- Elementos clave del aprendizaje reforzado
- Ejemplos

2. Métodos tabulares

- Multi-arm bandits
- Markov decision processes
- Programación Dinámica
- Métodos de Monte Carlo
- Temporal difference learning
- N-step learning

3. Métodos aproximados

- On-policy prediction
- On-policy control
- Off-policy methods
- Policy gradient methods

4. Deep reinforcement learning

- Deep Q-Networks
- Asynchronous Actor Critic
- Proximal Policy Optimization

5. Tópicos avanzados en aprendizaje reforzado

- Revisión del estado del arte de aprendizaje reforzado
- Revisión de las limitaciones actuales de aprendizaje reforzado

4. Metodología

- Clases (expositivas y prácticas)
- Lectura de papers
- Tareas teóricas y prácticas
- Proyecto de curso

5. Evaluación

Las evaluaciones del curso consistirán en tareas (**T**), controles (**C**) y un proyecto de curso (**P**). La nota final del curso será un promedio ponderado de estas evaluaciones:

$$\mathbf{NF} = 0,4 \cdot \mathbf{T} + 0,3 \cdot \mathbf{C} + 0,3 \cdot \mathbf{P}$$

Para aprobar el curso **NF** tiene que ser mayor o igual a 3,95. Todas las notas serán calculadas con **dos decimales**, salvo la nota final del curso que se calculará con **un decimal**.

El curso contará con 4 tareas. La nota de tareas será igual al promedio de esas 4 tareas:

$$\mathbf{T} = \frac{1}{4} (\mathbf{T}_1 + \mathbf{T}_2 + \mathbf{T}_3 + \mathbf{T}_4)$$

Al final de cada clase habrá un control. El control consistirá en preguntas relacionadas a la materia vista durante la clase o a un paper que deban leer antes de la clase. Cada respuesta correcta en un control valdrá un punto. Al final del semestre sumaremos todos los puntos obtenidos y los escalaremos para obtener la nota de los controles (**C**).

El proyecto del curso es en parejas y consiste en una propuesta de proyecto (**P**₁), una presentación de avance (**P**₂), una presentación final (**P**₃) y un informe final (**P**₄). La nota final del proyecto se calcula de la siguiente manera:

$$\mathbf{P} = \min \left(\frac{1}{4} (\mathbf{P}_1 + \mathbf{P}_2 + \mathbf{P}_3 + \mathbf{P}_4), \mathbf{P}_4 \right)$$

6. Fechas tentativas

Las fechas tentativas de las tareas y entregas del proyecto son las siguientes:

- [20/03 a las 23:59]: Tarea 1.
- [10/04 a las 23:59]: Tarea 2.
- [06/05 a las 23:59]: Tarea 3.
- [27/05 a las 23:59]: Tarea 4.
- [29/05 a las 23:59]: Propuesta proyecto.
- [16/06 al 20/06]: Presentaciones de avance proyecto.
- [01/07 al 04/07]: Presentaciones de proyecto.
- [07/07 a las 23:59]: Informe final proyecto.

7. Bibliografía

- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.
- François-Lavet, V., Henderson, P., Islam, R., Bellemare, M. G., & Pineau, J. (2018). *An introduction to deep reinforcement learning*. Foundations and Trends in Machine Learning, 11(3-4), 219-354.
- Ghavamzadeh, M., Mannor, S., Pineau, J., & Tamar, A. (2015). *Bayesian reinforcement learning: A survey*. Foundations and Trends in Machine Learning, 8(5-6), 359-483.
- Bellemare, M. G., Dabney, W., & Rowland, M. (2023). *Distributional reinforcement learning*. MIT Press.
- Proceedings de las siguientes conferencias: NeurIPS, ICML, ICLR, AAAI, IJCAI, AAMAS, entre otras.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.

Anexo: Política de integridad académica del Departamento de Ciencia de la Computación

Se espera los alumnos de la Escuela de Ingeniería de la Pontificia Universidad Católica de Chile mantengan altos estándares de honestidad académica, acorde al Código de Honor de la Universidad. Cualquier acto deshonesto o fraude académico está prohibido; los alumnos que incurran en este tipo de acciones se exponen a un Procedimiento Sumario. Es responsabilidad de cada alumno conocer y respetar el documento sobre Integridad Académica publicado por la Dirección de Pregrado de la Escuela de Ingeniería (Disponible en SIDING, en la sección Pregrado/Asuntos Estudiantiles/Reglamentos/Reglamentos en Ingeniería/Integridad Académica).

Específicamente, para los cursos del Departamento de Ciencia de la Computación, rige obligatoriamente la siguiente *política de integridad académica*. Todo trabajo presentado por un alumno para los efectos de la evaluación de un curso debe ser hecho **individualmente** por el alumno, **sin apoyo en material de terceros**. Por “trabajo” se entiende en general las interrogaciones escritas, las tareas de programación u otras, los trabajos de laboratorio, los proyectos, el examen, entre otros.

En particular, si un alumno copia un trabajo, o si a un alumno se le prueba que compró o intentó comprar un trabajo, **obtendrá nota final 1.1 en el curso** y se solicitará a la Dirección de Pregrado de la Escuela de Ingeniería que no le permita retirar el curso de la carga académica semestral.

Por “copia” se entiende incluir en el trabajo presentado como propio, partes hechas por otra persona. En caso que corresponda a “copia” a otros alumnos, la sanción anterior se aplicará a todos los involucrados. En todos los casos, se informará a la Dirección de Pregrado de la Escuela de Ingeniería para que tome sanciones adicionales si lo estima conveniente.

Obviamente, está permitido usar material disponible públicamente, por ejemplo, libros o contenidos tomados de Internet, **siempre y cuando se incluya la referencia correspondiente.**

Lo anterior se entiende como complemento al Reglamento del Alumno de la Pontificia Universidad Católica de Chile (<http://admisionyregistros.uc.cl/alumnos/informacion-academica/reglamentos-estudiantiles>). Por ello, es posible pedir a la Universidad la aplicación de sanciones adicionales especificadas en dicho reglamento.