



N      20/2011      199

(4)  
cuatro

**Facultad de Ingeniería**  
**Comisión Académica de Posgrado**

**Formulario de Aprobación Curso de Posgrado**

**Asignatura:** Aprendizaje Profundo por Refuerzo

(Si el nombre contiene siglas deberán ser aclaradas)

**Profesor de la asignatura**<sup>1</sup>: Dr. Pablo Sprechmann - DeepMind - Reino Unido.  
(título, nombre, grado o cargo, Instituto o Institución)

**Profesor Responsable Local**<sup>1</sup>: Dr. Mauricio Delbracio - Gr. 3 - IIE.  
(título, nombre, grado, Instituto)

**Otros docentes de la Facultad:** Dr. José Lezama - Gr. 3 - IIE.  
(título, nombre, grado, Instituto)

**Docentes fuera de Facultad:**  
(título, nombre, cargo, Institución, país)

**Programa(s):** Maestría y Doctorado en Ingeniería Eléctrica

**Instituto ó Unidad:** IIE

**Departamento ó Area:** Procesamiento de señales, Aprendizaje Automático

<sup>1</sup> 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)

**Horas Presenciales:** 20

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

**Nº de Créditos:** 5

(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 la enseñanza)

**Público objetivo y Cupos:** Estudiantes de Maestría en Ingeniería Matemática, Maestría y Doctorado en Ingeniería Eléctrica, Estadística, Computación y carreras afines. No hay cupos.

**Objetivos:**

El objetivo del curso es presentar una introducción a *reinforcement learning* (aprendizaje por refuerzos) y al *deep reinforcement learning*, de forma que los estudiantes sean capaces de implementar, aplicar y evaluar algoritmos relevantes utilizando la herramienta Tensorflow. El *reinforcement learning* es un área del aprendizaje de máquinas cuyo objetivo es determinar qué acciones debe escoger un agente en un entorno dado con el fin de maximizar alguna noción de "recompensa" o premio acumulado. El *deep reinforcement learning* es una sub-área del *reinforcement learning* en donde redes neuronales (entrenadas mediante técnicas de aprendizaje profundo) son utilizadas como funciones aproximantes por algoritmos de *reinforcement learning*.

El aprendizaje profundo por refuerzo tiene potencialmente una gran cantidad de aplicaciones. Cualquier aplicación en la que necesite encontrar la mejor política para tomar una acción con el fin de maximizar un objetivo dado en contexto complejo (difícil de modelar) es un buen candidato de aplicación. Algunos ejemplos son: entrenamiento de robots o tareas robóticas, control del tráfico urbano, administración de sistemas de energía con fuentes y configuraciones de generación múltiples, piloto automático de vehículos autónomos, comercio con estrategias óptimas.

El curso cubre dos partes. La primera es una introducción a los problemas de predicción y control utilizando técnicas del *reinforcement learning*. La segunda, aborda el *deep reinforcement learning*, en particular los recientes avances

en el área. El objetivo es cubrir los aspectos teóricos básicos del aprendizaje por refuerzos, y los principales desarrollos algorítmicos que han aparecido en los últimos años en el área de *deep reinforcement learning*. El curso busca presentar a los estudiantes los principales aspectos de modelado, algorítmicos y de optimización de forma de que ellos mismos sean capaces de implementar sus propios modelos.

**Conocimientos previos exigidos:** Conocimientos básicos de cálculo diferencial, álgebra lineal, probabilidad y programación (conocimientos sólidos en al menos un lenguaje de programación).

**Conocimientos previos recomendados:** Familiaridad con conceptos básicos de reconocimiento de patrones, procesamiento de señales, programación en lenguaje Python y Optimización.

**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)

- Horas clase (teórico): 15
- Horas clase (práctico): 0
- Horas clase (laboratorio):
- Horas consulta: 5
- Horas evaluación: 0
  - Subtotal horas presenciales: 20
- Horas estudio: 20
- Horas resolución ejercicios/prácticos: 35
- Horas proyecto final/monografía: 0
  - Total de horas de dedicación del estudiante: 75

**Forma de evaluación:** La evaluación consiste en la entrega de una serie de ejercicios obligatorios en máquina y la entrega de respuestas a preguntas teóricas.

**Temario:**

Se intentarán cubrir los siguientes temas: procesos de decisión Markovianos, planeamiento via programación dinámica, *model-free prediction and control*, *policy gradient methods* y el dilema de la explotación/exploración. Posibles aplicaciones a implementar incluyen: entrenar agentes para que sean capaces de aprender a jugar juegos de mesa y videojuegos.

Tentativamente,

- 1) Introducción a los paradigmas de *reinforcement learning*, repasando los principales fundamentos teóricos.
- 2) Introducción al *deep reinforcement learning*. (DQN, *deep policy gradients*, A3C)
- 3) Introducción a las técnicas básicas de optimización para redes neuronales y su aplicación al *reinforcement learning*.
- 4) Breve introducción a Tensorflow.
- 5) Aplicación de algoritmos de *reinforcement learning* a entornos con dinámicas complejas

**Bibliografía:**

La bibliografía seguirá varios artículos y algunos capítulos de libros publicados recientemente.

- [1] Sutton, Richard S., & Andrew G. Barto. *Reinforcement learning: An introduction*. 2nd edition, MIT press, 2018.
- [2] Goodfellow, I., Bengio, Y., Courville, A., & Bengio, *Deep learning*. MIT press, 2016
- [3] Szepesvári, C., *Algorithms for reinforcement learning. Synthesis lectures on artificial intelligence and machine learning*, Morgan & Claypool, 2010
- [4] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G. and Petersen, S., 2015. Human-level control through deep reinforcement learning. *Nature*, 518(7540), p.529.
- [5] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M. and Kudlur, M., 2016, November. *Tensorflow: a system for large-scale machine learning*. In *OSDI (Vol. 16, pp. 265-283)*.
- [6] LeCun, Y., Bengio, Y. and Hinton, G., 2015. *Deep learning*. *nature*, 521(7553), p.436.
- [7] Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M. and Dieleman, S., 2016. *Mastering the game of Go with deep neural networks and tree search*. *nature*, 529(7587), p.484.



Facultad de Ingeniería  
Comisión Académica de Posgrado

Datos del curso

Fecha de inicio y finalización: del 25 de febrero al 1 de marzo de 2019

Horario y Salón: A confirmar.