

### Formulario de Aprobación Curso de Posgrado

**Asignatura:** Aprendizaje Profundo para Visión Artificial

(Si el nombre contiene siglas deberán ser aclaradas)

---

**Profesor de la asignatura**<sup>1</sup>: Dr. Mauricio Delbracio – Gr. 3 – IIE.

(título, nombre, grado o cargo, Instituto o Institución)

**Profesor Responsable Local**<sup>1</sup>:

(título, nombre, grado, Instituto)

**Otros docentes de la Facultad:**

Dr. José Lezama – Investigador Postdoctoral – IIE

Msc. Guillermo Carbajal – Gr. 2 – IIE

(título, nombre, grado, Instituto)

**Docentes fuera de Facultad:**

(título, nombre, cargo, Institución, país)

**Instituto ó Unidad:** IIE

**Departamento ó Área:** Procesamiento de Señales

<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:** 48

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

**Nº de Créditos:** 8

(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:** Este curso busca introducir las redes neuronales profundas y su aplicación a la visión artificial a estudiantes de posgrado en Ingeniería Eléctrica, Ingeniería en Computación o Matemática. No hay cupos.

(si corresponde, se indicará el número de plazas, mínimo y máximo y los criterios de selección. Asimismo, se adjuntará en nota aparte los fundamentos de los cupos propuestos. Si no existe indicación particular para el cupo máximo, el criterio general será el orden de inscripción en el Depto. de Posgrado, hasta completar el cupo asignado)

---

**Objetivos:**

En los últimos años, el aprendizaje profundo (*deep learning*) se ha convertido en una herramienta fundamental en el aprendizaje de máquinas para una amplia variedad de dominios y aplicaciones. Uno de sus mayores éxitos ha sido su aplicación a la visión artificial, donde el desempeño en problemas como el reconocimiento de objetos y acciones ha permitido importantes mejoras en la última década.

El aprendizaje profundo es una técnica que emplea redes neuronales para aprender representaciones a partir de una serie de datos observados, que puedan ser de utilidad para resolver problemas de alto nivel como ser predicción, restauración o clasificación de señales. Dichas representaciones se denominan profundas por estar construidas a partir de una jerarquía compuesta de sucesivas capas que representan las observaciones con un nivel creciente de abstracción. El aprendizaje profundo pretende descubrir las propiedades intrínsecas de grandes

volúmenes de datos construyendo representaciones distribuidas, tanto en contextos supervisados como no supervisados.

Este curso brinda una introducción al aprendizaje profundo con un enfoque práctico en la visión artificial. El objetivo principal es presentar a los estudiantes los principales aspectos de modelado, algorítmicos y de optimización de forma de que ellos mismos puedan implementar (diseñar, entrenar y validar) sus propios modelos. Se presentarán modelos simples basados en redes convolucionales que permiten explicar matemáticamente algunas de las propiedades claves para su éxito. Todos los conceptos serán ilustrados con aplicaciones específicas en los campos de visión artificial y procesamiento de imágenes.

---

**Conocimientos previos exigidos:** Conocimientos básicos de Cálculo diferencial, Álgebra Lineal, Probabilidad y Estadística y Programación (conocimientos sólidos en al menos un lenguaje de programación).

**Conocimientos previos recomendados:** Familiaridad con conceptos básicos de reconocimiento de patrones, procesamiento de imágenes, optimización, programación en lenguaje Python.

---

**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)

Se darán 36 hs. de clases teórico-prácticas divididas en 24 clases de una hora y media cada una, a razón de dos clases por semana. Se estima una dedicación de 20 horas por parte del estudiante para estudiar los temas impartidos en clase, 36 horas para resolver los ejercicios de práctico y 16 para la realización del trabajo final.

- Horas clase (teórico): 27
- Horas clase (práctico): 9
- Horas clase (laboratorio): 0
- Horas consulta: 8
- Horas evaluación: 4
  - Subtotal horas presenciales: 48
- Horas estudio: 20
- Horas resolución ejercicios/prácticos: 36
- Horas proyecto final/monografía: 16
  - Total de horas de dedicación del estudiante: 120

---

**Forma de evaluación:**

La evaluación consiste en la entrega de una serie de ejercicios obligatorios en máquina, la entrega de respuestas a preguntas teóricas y la realización de un mini-proyecto final con defensa oral. El mini-proyecto final consistirá en el estudio de un trabajo publicado recientemente o bien en evaluar y profundizar en alguna de las técnicas discutidas durante el curso.

**Temario:** Tentativamente, se intentarán cubrir los siguientes temas

1. Introducción al aprendizaje profundo, motivación, reseña histórica. (1,5h)
2. Aprendizaje supervisado. Métodos lineales de clasificación, regla del k-vecino más cercano. (3h)
3. Representaciones de alto nivel, características en imágenes. (1,5h)
4. Formulación del aprendizaje como un problema de optimización. Algoritmo de "backpropagation". (3h)
5. Redes neuronales totalmente conectadas y su entrenamiento. (6h)
6. Redes neuronales de convolución, arquitectura, capas, operadores, y su entrenamiento. (6h)
7. Análisis de redes neuronales, visualización de representaciones. Propiedades: invarianza, covarianza, redundancia e invertibilidad. (3h)
8. Redes neuronales recurrentes (RNN), redes de corta-larga memoria (LSTM). Entrenamiento y ejemplos. (3h)
9. Redes adversas generativas (GAN). Entrenamiento y ejemplos. (3h)
10. Aprendizaje profundo en la práctica (bibliotecas). Aplicaciones (clasificación y subtitulado de imágenes, predicción de video, superresolución, eliminación de ruido y borrosidad). (6h)

Las horas indicadas son la combinación de clases teóricas y laboratorios en máquina.

---

### Bibliografía:

- Goodfellow, I., Bengio, Y., Courville, A. *"Deep learning"*. MIT Press, 2016.
- Hastie, T, Tibshirani, R. Friedman, J. *"The Elements of Statistical Learning"*. NY Springer, 2001
- LeCun, Y., Bengio, Y., Hinton, G. *"Deep learning"*. Nature 521.7553 (2015): 436-444.
- Mallat, S., *"Understanding deep convolutional networks"* Phil. Trans. R. Soc. A 374.2065 (2016): 20150203.
- Krizhevsky, A., Sutskever, I., Hinton, G. *"Imagenet classification with deep convolutional neural networks"*. NIPS, 2012.
- Zeiler, M. D., and Fergus, R. *"Visualizing and understanding convolutional networks"*. ECCV, 2014.
- Li, Fei-Fei, L., Karpathy, A., Johnson, J. *"CS231n: Convolutional neural networks for visual recognition"*. Stanford University, notas de curso, 2015.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. *"Generative adversarial nets"*. NIPS, 2014
- Anh, N., Yosinski, J., Clune, J. *"Deep neural networks are easily fooled: High confidence predictions for unrecognizable images."* CVPR, 2015
- Johnson, J., Alahi, A., Fei-Fei, L. *"Perceptual losses for real-time style transfer and super-resolution"*. ECCV 2016.

**Datos del curso**

**Fecha de inicio y finalización:** segundo semestre

**Horario y Salón:** A confirmar.