

Formulario de aprobación de curso de posgrado/educación permanente

Asignatura: Redes Neuronales para Lenguaje Natural

(Si el nombre contiene siglas deberán ser aclaradas)

Modalidad: (posgrado, educación permanente o ambas)	Posgrado	X
	Educación permanente	

Profesor de la asignatura ¹:

Profesor Adjunto Dr. Ing. Luis Chiruzzo, Instituto de Computación

Profesor Responsable Local ¹:

Otros docentes de la Facultad:

Asistente MSc. Ing. Mathias Etcheverry, Instituto de Computación

Profesor Adjunto Dr. Ing. Guillermo Moncecchi, Instituto de Computación

Docentes fuera de Facultad:

(título, nombre, cargo, institución, país)

¹ 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)

[Si es curso de posgrado]

Programa(s) de posgrado: Maestría / Doctorado en Informática Pedeciba-Udelar, Maestría en Ciencia de Datos y Aprendizaje Automático (CDAA)

Instituto o unidad: Instituto de Computación

Departamento o área: Procesamiento de Lenguaje Natural

Horas Presenciales: 55

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

Nº de Créditos: 10

[Exclusivamente para curso de posgrado]

(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 enseñanza)

Público objetivo: Estudiantes de posgrado en Informática o Ciencia de Datos.

Cupos: Sin cupo.

(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, hasta completar el cupo asignado)

Objetivos: El objetivo del curso es presentar los fundamentos y las técnicas asociados a los modelos de redes neuronales artificiales, con énfasis en su aplicación al procesamiento de lenguaje natural. El

estudiante obtiene durante el curso el conocimiento de base sobre diferentes modelos de redes neuronales, así como la capacidad de aplicarlos a problemas concretos, utilizando bibliotecas de código abierto.

Conocimientos previos exigidos: Programación, Álgebra Lineal, Cálculo Diferencial.

Conocimientos previos recomendados: Conocimientos de Aprendizaje Automático, Probabilidad y Estadística.

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)

Descripción de la metodología:

[Obligatorio]

El curso está organizado en unidades temáticas. Se dictarán dos clases semanales en modalidad teórico/práctica y los estudiantes realizarán trabajos de laboratorio en grupo.

Detalle de horas:

- Horas de clase (teórico/práctico): 40
- Horas de consulta: 12
- Horas de evaluación: 3
 - Subtotal de horas presenciales: 55
- Horas de estudio: 35
- Horas de resolución de ejercicios/prácticos: 30
- Horas proyecto final/monografía: 30
 - Total de horas de dedicación del estudiante: 150

Forma de evaluación:

El curso se evalúa en base al laboratorio y la prueba final, contando cada una con 50% del puntaje del curso. Para cada instancia (laboratorio y prueba final), se debe obtener como mínimo un 60% de los puntos. La prueba final se realiza de manera individual.

Temario:

1- Introducción a redes neuronales

Regresión logística. Unidad (neurona). Función de activación. Aprendizaje: funciones de pérdida, métodos de descenso por gradiente. Perceptron multicapa (redes feedforward). Algoritmo de backpropagation. Representaciones básicas de texto (bag of words, tf-idf).

2- Representaciones distribuidas de palabras: Word Embeddings

Hipótesis distribucional, Word embeddings, uso como input para tareas de PLN. Representación de unidades más pequeñas que la palabra.

3- Arquitecturas clásicas de aprendizaje profundo

Modelos convolucionales, modelos recurrentes, LSTM.

4- Modelo atencional

Definición del modelo atencional, uso en modelos recurrentes.

5- Modelo transformer

Modelo auto-atencional, arquitectura transformer, BERT.

6- Grandes modelos de lenguaje (LLMs)

Arquitectura transformer utilizada para modelos más complejos (large language models - LLMs) como GPT o T5. Uso para resolución de problemas de PLN.

Bibliografía:

(título del libro-nombre del autor-editorial-ISBN-fecha de edición)

Speech and Language Processing (Third ed.). Jurafsky, D. and J. H. Martin (2021).
<https://web.stanford.edu/~jurafsky/slp3/>

Neural Network Methods in Natural Language Processing (Synthesis Lectures on Human Language Technologies, 37) - Yoav Goldberg - ISBN 1627052984 - Morgan & Claypool Publishers (2017)

Deep Learning - Ian Goodfellow, Yoshua Bengio, Aaron Courville - MIT Press (2016)

Para cada unidad temática se indicará bibliografía específica adicional.
