



Programa de REDES NEURONALES PARA LENGUAJE NATURAL

1. NOMBRE DE LA UNIDAD CURRICULAR

Redes Neuronales para Lenguaje Natural

2. CRÉDITOS

10 créditos

3. OBJETIVOS DE LA UNIDAD CURRICULAR

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.

4. METODOLOGÍA DE ENSEÑANZA

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. Se realizará una prueba escrita individual final para evaluar los conocimientos obtenidos en el curso.

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

5. TEMARIO

Incluye una descripción general de los grandes temas del curso y de los subtemas incluidos en cada uno de ellos.

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.

6. BIBLIOGRAFÍA

Identificación de las publicaciones básicas y complementarias adecuadas para el buen seguimiento del curso. Se debería observar la disponibilidad de estos textos, tanto en la Biblioteca de Facultad como en el mercado. En caso de existir varios textos principales, indicar para qué tema aporta cada uno. La referencia bibliográfica deberá darse de la siguiente forma:

Tema	Básica	Complementaria
Introducción a redes neuronales	(1)	(2, 3)
Representaciones distribuidas de palabras	(1)	(2, 3)
Arquitecturas clásicas de aprendizaje profundo	(1)	(2, 3)
Modelo atencional	(1)	(2)
Modelo transformer	(1)	
Grandes modelos de lenguaje	(1)	

6.1 Básica

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

6.2 Complementaria

2. Neural Network Methods in Natural Language Processing (Synthesis Lectures on Human Language Technologies, 37) - Yoav Goldberg - ISBN 1627052984 - Morgan & Claypool Publishers (2017)
3. Deep Learning - Ian Goodfellow, Yoshua Bengio, Aaron Courville - MIT Press (2016)

7. CONOCIMIENTOS PREVIOS EXIGIDOS Y RECOMENDADOS

7.1 Conocimientos Previos Exigidos: Se debe contar con formación en Álgebra, Probabilidad y Estadística y Aprendizaje Automático.

7.2 Conocimientos Previos Recomendados: Es de interés contar con formación en Procesamiento de Lenguaje Natural.

ANEXO A
Para todas las Carreras

A1) INSTITUTO

Instituto de Computación

A2) CRONOGRAMA TENTATIVO

Semana 1	Tema 1: Introducción a redes neuronales 3 horas teórico/práctico
Semana 2	Tema 1: Introducción a redes neuronales 3 horas teórico/práctico 1 hora consulta
Semana 3	Tema 2: Representaciones distribuidas de palabras 3 horas teórico/práctico
Semana 4	Tema 2: Representaciones distribuidas de palabras 3 horas teórico/práctico 1 hora consulta
Semana 5	Tema 3: Arquitecturas clásicas de aprendizaje profundo 3 horas teórico/práctico
Semana 6	Tema 3: Arquitecturas clásicas de aprendizaje profundo 3 horas teórico/práctico
Semana 7	Tema 3: Arquitecturas clásicas de aprendizaje profundo 3 horas teórico/práctico 1 hora consulta
Semana 8	Tema 4: Modelo atencional 3 horas teórico/práctico 1 hora consulta
Semana 9	Tema 5: Modelo transformer 3 horas teórico/práctico
Semana 10	Tema 5: Modelo transformer 3 horas teórico/práctico
Semana 11	Tema 5: Modelo transformer 3 horas teórico/práctico 1 hora consulta
Semana 12	Tema 6: Grandes modelos de lenguaje 3 horas teórico/práctico
Semana 13	Tema 6: Grandes modelos de lenguaje 4 horas teórico/práctico 1 hora consulta
Semana 14	3 horas consulta
Semana 15	3 horas consulta
Segundos parciales	Prueba escrita: 3 horas

A3) MODALIDAD DEL CURSO Y PROCEDIMIENTO DE EVALUACIÓN

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. Se realizará una prueba escrita individual final para evaluar los conocimientos obtenidos en el curso.

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.

A4) CALIDAD DE LIBRE

No se puede acceder a la calidad de libre.

A5) CUPOS DE LA UNIDAD CURRICULAR

No tiene

ANEXO B para la(s) carrera(s) Ingeniería en Computación (plan 97)/ Licenciatura en Computación

B1) ÁREA DE FORMACIÓN

Inteligencia Artificial y Robótica

B2) UNIDADES CURRICULARES PREVIAS

Para el Curso, examen de:

Aprendizaje Automático
Probabilidad y Estadística
Programación 3

Para el Examen:

No aplica (no hay exámenes).

APROB RES. CONSEJO DE FAC. 1113.
Fecha 16/5/23 Exp. 060120-00034-23