



FICHA TÉCNICA DE LA ASIGNATURA

Datos de la asignatura	
Nombre completo	Procesamiento de Lenguaje Natural II
Código	DTC-IMAT-416
Título	Grado en Ingeniería Matemática e Inteligencia Artificial [GiMAT 25]
Impartido en	Grado en Ingeniería Matemática e Inteligencia Artificial [Cuarto Curso]
Créditos	6,0 ECTS
Carácter	Obligatoria (Grado)
Departamento / Área	Departamento de Ciencias de la Computación e Inteligencia Artificial

Datos del profesorado	
Profesor	
Nombre	Andrés Occhipinti Liberman
Departamento / Área	Departamento de Ciencias de la Computación e Inteligencia Artificial (DCIA)
Correo electrónico	aocchipinti@icai.comillas.edu
Profesor	
Nombre	Sergio Romero Tapiador
Departamento / Área	Departamento de Ciencias de la Computación e Inteligencia Artificial (DCIA)
Correo electrónico	sromero@comillas.edu

DATOS ESPECÍFICOS DE LA ASIGNATURA

Contextualización de la asignatura
Aportación al perfil profesional de la titulación
Este curso es una continuación de <i>Procesamiento del Lenguaje Natural I</i>. Se centra en el estudio de las arquitecturas modernas de procesamiento del lenguaje natural (PLN), con especial énfasis en el modelo transformer. El objetivo es proporcionar una base sólida en la teoría y las técnicas utilizadas para llevar a cabo diferentes tareas de PLN, como la comprensión y la generación de lenguaje natural. El curso introduce técnicas de PLN aplicadas tanto al texto como al habla, así como conceptos de web semántica, ontologías y grafos de conocimiento.
Prerrequisitos
Haber cursado una asignatura de Aprendizaje Automático y de introducción al Procesamiento del Lenguaje Natural.

Resultados del proceso de formación y aprendizaje
Conocimientos o contenidos



CON02	Conocimientos básicos sobre el uso y programación de los ordenadores, sistemas operativos, bases de datos y programas informáticos con aplicación en ingeniería.
CON13	Conocimiento y capacidad para utilizar distintas tecnologías de procesamiento, representación y análisis de lenguaje natural.
CON17	Capacidad para analizar el comportamiento de los sistemas cognitivos y aplicarlos en el dominio artificial. Conocimiento de los principios de los procesos psicológicos básicos.
Habilidades o destrezas	
HAB03	Capacidad para identificar, analizar y definir los elementos significativos que constituyen un problema vinculado a la explotación de datos e inteligencia artificial aplicada a las actividades empresariales para resolverlo con criterio y de forma efectiva.
HAB23	Capacidad para especificar, diseñar e implementar las técnicas de aprendizaje automático y profundo para la resolución de problemas complejos.

BLOQUES TEMÁTICOS Y CONTENIDOS

Contenidos – Bloques Temáticos
1. Arquitecturas de Transformers. 2. Entrenamiento de modelos de lenguaje a gran escala (LLMs): Pre-entrenamiento y post-entrenamiento (fine-tuning y preference tuning). 3. Transformers eficientes: fine-tuning eficiente en parámetros; técnicas de comprensión de modelos; cuantización. 4. Aplicaciones de transformers. 5. Transformers aplicados al procesamiento del habla. 6. Web semántica, ontologías, y grafos de conocimiento.



METODOLOGÍA DOCENTE

Aspectos metodológicos generales de la asignatura
Metodología Presencial: Actividades
Las actividades formativas serán: Clases magistrales expositivas y participativas: - El profesor combinará exposición de los contenidos teóricos con ejemplos prácticos, tanto matemáticos como de programación. - El alumno dispondrá de algunos ejemplos prácticos, generados dentro y fuera del aula por el profesor. Sesiones prácticas, ejercicios prácticos y resolución de problemas: - El alumno resolverá problemas relacionados con el contenido teórico presentado anteriormente por el profesor, fomentando la absorción del contenido teórico y las dinámicas de trabajo cooperativo. - Se considerará como posible actividad que los alumnos (individualmente o en grupos) presenten su resolución de ejercicios en clase y se trabajará sobre dicha resolución para mejorarla o discutir detalles de la misma. - Las prácticas se graduarán por dificultad que los alumnos irán implementando conforme vayan finalizando cada hito. - Se incluirán también ejercicios de cálculo a mano. - Se considerará realizar pruebas, desarrollar prácticas complementarias a las semanales y retos gamificados.
Metodología No presencial: Actividades
Las actividades formativas serán: Ejercicios prácticos y resolución de problemas: El alumno dispondrá de problemas concretos enfocados a asimilar los conceptos explicados teóricos en la sesión anterior de teoría para desarrollar de forma no presencial. Estudio personal: El objetivo principal del trabajo no presencial es llegar a entender y comprender los conceptos teóricos de la asignatura, así como ser capaz de poner en práctica estos conocimientos para resolver los diferentes tipos de problemas.

RESUMEN HORAS DE TRABAJO DEL ALUMNO

HORAS PRESENCIALES		
AF3. Sesiones prácticas con uso de software	AF5. Tutorías para resolución de dudas	
20.00	5.00	
HORAS NO PRESENCIALES		
AF3. Sesiones prácticas con uso de software	AF10. Estudio personal	AF6. Trabajos
30.00	55.00	30.00
CRÉDITOS ECTS: 6,0 (140,00 horas)		



EVALUACIÓN Y CRITERIOS DE CALIFICACIÓN

El uso de IA para crear trabajos completos o partes relevantes, sin citar la fuente o la herramienta o sin estar permitido expresamente en la descripción del trabajo, será considerado plagio y regulado conforme al Reglamento General de la Universidad.

Actividades de evaluación	Criterios de evaluación	Peso
Trabajo final	Presentación oral y evaluación de la entrega.	15
Evaluación continua	Trabajos de evaluación continua (entregables)	10
Exámenes.	- Examen intersemestral: 30% - Examen Final: 45%	75

Calificaciones

La calificación final en convocatoria ordinaria y extraordinaria de la asignatura dependerá de la evaluación de las siguientes actividades:

Elemento evaluable	Porcentaje
Examen final	45%
Examen intersemestral	30%
Proyecto final	15%
Trabajos de evaluación continua (entregables)	10%

Para aprobar la asignatura en la convocatoria ordinaria, los alumnos tienen que obtener al menos 5 puntos sobre 10 en el examen final de la asignatura y en el proyecto final.

En el caso de acudir a la convocatoria extraordinaria, para aprobar la asignatura, se deberán obtener al menos 5 puntos sobre 10 en el examen extraordinario. En esta convocatoria, se calcularán dos medias:

- Media 1: El examen extraordinario vale un 75%
- Media 2: El examen extraordinario vale un 45%, y el examen intersemestral vale un 30%

Se aplicará la media más favorable para el alumno.

La inasistencia al 15% o más de las horas presenciales de esta asignatura puede tener como consecuencia la imposibilidad de presentarse a las convocatorias ordinaria y extraordinaria. La práctica final será grupal.

Uso de IA generativa:

Es fundamental que intentes realizar los ejercicios de cálculo y programación de manera autónoma (sin IA), porque así podrás asimilar mejor el contenido y llegar más preparado a los exámenes. Puedes apoyarte en la IA para planificar tu trabajo, generar ideas iniciales o resolver dudas puntuales. Eso sí, recuerda evaluar siempre de forma crítica la información que te proporcione, ya que no siempre es fiable.



BIBLIOGRAFÍA Y RECURSOS

Bibliografía Básica

Jurafsky, D., & Martin, J. H. (25). Speech and language processing (tercera edición, borrador). Disponible en: <https://web.stanford.edu/~jurafsky/slp3/>

Vaswani, A. "Attention is all you need." *Advances in Neural Information Processing Systems* (2017).