



FICHA TÉCNICA DE LA ASIGNATURA

Datos de la asignatura	
Nombre completo	Procesamiento de lenguaje natural I
Código	DTC-IMAT-323
Título	Grado en Ingeniería Matemática e Inteligencia Artificial
Impartido en	Grado en Ingeniería Matemática e Inteligencia Artificial [Tercer Curso]
Créditos	6,0 ECTS
Carácter	Obligatoria (Grado)
Departamento / Área	Departamento de Telemática y Computación
Responsable	Andrés Occhipinti Liberman

Datos del profesorado	
Profesor	
Nombre	Andrés Occhipinti Liberman
Departamento / Área	Departamento de Telemática y Computación
Correo electrónico	aocchipinti@icai.comillas.edu
Profesores de laboratorio	
Profesor	
Nombre	Jaime Pizarroso Gonzalo
Departamento / Área	Departamento de Telemática y Computación
Despacho	Santa Cruz de Marcenado 26
Correo electrónico	jpizarroso@comillas.edu
Teléfono	2732

DATOS ESPECÍFICOS DE LA ASIGNATURA

Contextualización de la asignatura
Aportación al perfil profesional de la titulación
Este curso es una introducción al Procesamiento del Lenguaje Natural (PLN). 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 abarca una combinación de técnicas, incluidos métodos basados en reglas, estadísticos y de aprendizaje automático.
Prerrequisitos
Haber cursado una asignatura de Aprendizaje Automático.
Competencias - Objetivos



Competencias

GENERALES

CG04	Conocimientos básicos sobre el uso y programación de los ordenadores, sistemas operativos, bases de datos y programas informáticos con aplicación en ingeniería.
CG08	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

ESPECÍFICAS

CE03	Capacidad para saber aplicar las técnicas matemáticas más adecuadas en la resolución de los diferentes problemas, técnicos y tecnológicos, planteados en el ámbito de la ingeniería y la inteligencia artificial. Aptitud para conocer el rango de aplicabilidad y limitaciones en la resolución de problemas de las diferentes herramientas matemáticas.
CE28	Conocimiento y capacidad para utilizar distintas tecnologías de procesamiento, representación y análisis de lenguaje natural.
CE31	Capacidad para especificar, diseñar e implementar las técnicas de aprendizaje automático y profundo para la resolución de problemas complejos.
CE36	Capacidad para analizar el comportamiento de los sistemas cognitivos y aplicarlos en el dominio artificial. Conocimiento de los principios de los procesos psicológicos básicos

Resultados de Aprendizaje

RA1	Conocer y distinguir los niveles del lenguaje natural y la complejidad de cada uno
RA2	Conocer las técnicas y fundamentos aplicados de procesado del lenguaje natural en cada uno de sus niveles
RA3	Conocer las herramientas básicas de procesado de la información para su posterior transformación
RA4	Aplicar los conocimientos aprendidos en otras materias como aprendizaje automático a problemas reales de generación de contenidos
RA5	Estar familiarizado con las herramientas y técnicas que se utilizan dentro de este ámbito

BLOQUES TEMÁTICOS Y CONTENIDOS

Contenidos – Bloques Temáticos

1. Principios básicos del procesamiento del lenguaje natural.
2. Clasificación textual. Clasificadores probabilísticos y neurales.
3. Modelos de lenguaje. El modelo n-grama y modelos neurales (feed-forward y recurrentes). Evaluación de modelos de lenguaje.
4. Word embeddings y semántica de vectores.
5. Análisis de estructura lingüística. Dependency parsing y constituency parsing.
6. Representación lógica de la semántica del lenguaje natural.



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 y con ejemplos prácticos, tanto matemáticos como de programación.
- El alumno dispondrá de algunos ejemplos prácticos de código, generado dentro y fuera del aula por el profesor.

Ejercicios prácticos y resolución de problemas:

- En la segunda clase semanal, 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.
- Puntualmente, alumnos (individualmente o en grupos) presentarán su resolución de ejercicios en clase y se trabajará sobre dicha resolución para mejorarla o discutir detalles de la misma.

Sesiones prácticas con uso de software:

- Se dedicarán las sesiones prácticas a realizar una práctica semanal, dirigida a consolidar los contenidos teóricos presentados previamente.
- Las prácticas se graduarán por dificultad que los alumnos irán implementando conforme vayan finalizando cada hito.

Actividades de evaluación continua del rendimiento:

- Se realizarán pruebas, desarrollarán prácticas complementarias a las semanales y retos gamificados.

CG04, CG08, CE03, CE28,
CE31, CE36

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

CG04, CG08, CE03, CE28,
CE31, CE36



los diferentes tipos de problemas.

RESUMEN HORAS DE TRABAJO DEL ALUMNO

HORAS PRESENCIALES				
Clases magistrales expositivas y participativas	Sesiones prácticas con uso de software	Tutorías para resolución de dudas	Ejercicios prácticos y resolución de problemas	Actividades de evaluación continua del rendimiento
28.00	20.00	5.00	10.00	2.00
HORAS NO PRESENCIALES				
Sesiones prácticas con uso de software	Estudio personal	Trabajos		
30.00	55.00	30.00		
CRÉDITOS ECTS: 6,0 (180,00 horas)				

EVALUACIÓN Y CRITERIOS DE CALIFICACIÓN

Actividades de evaluación	Criterios de evaluación	Peso
Exámenes escritos evaluando conceptos de la asignatura.	<ul style="list-style-type: none">Examen intersemestral: 30%Examen Final: 45%	75 %
Trabajo final	<ul style="list-style-type: none">Tendrá un peso de un 15% de la calificación.	15 %
Evaluación continua	Trabajos de evaluación continua (entregables): <ol style="list-style-type: none">Clasificación textualModelos de lenguajeAnálisis sintáctico y semántico	10 %

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):

1. Clasificación textual
2. Modelos de lenguaje
3. Análisis sintáctico y semántico

10%

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

BIBLIOGRAFÍA Y RECURSOS

Bibliografía Básica

- Jurafsky, D., & Martin, J. H. (2008). *Speech and language processing (2nd ed.)*. Upper Saddle River, NJ: Pearson.
- Eisenstein J. *Introduction to Natural Language Processing*. Cambridge Massachusetts: MIT Press; 2019.

En cumplimiento de la normativa vigente en materia de **protección de datos de carácter personal**, le informamos y recordamos que puede consultar los aspectos relativos a privacidad y protección de datos que ha aceptado en su matrícula entrando en esta web y pulsando "descargar"

<https://servicios.upcomillas.es/sedelectronica/inicio.aspx?csv=02E4557CAA66F4A81663AD10CED66792>