UAB
Universitat Autònoma de Barcelona

# Procesado del Lenguaje Natural

2026/2027

Código: 104422
Créditos: 6

| Titulación | Tipo | Curso |
|---|---|---|
| Matemática Computacional y Analítica de Datos | OP | 4 |

## Profesor/a de contacto

Nombre : Ernest Valveny Llobet

Correo electrónico : ernest.valveny@uab.cat

## Equipo docente

Pau Torras Coloma

## Idiomas de los grupos

Puede consultar esta información al final del documento.

## Prerrequisitos

No existen requisitos previos oficiales pero se recomienda haber cursado las asignaturas de Fundamentos de Programación I y II, Fundamentos de Matemáticas I y II, Probabilidad y Estadística, Ingeniería de Datos y Fundamentos de Aprendizaje Automático.

## Objetivos

Este curso proporciona una descripción general de las técnicas fundamentales para el procesamiento del lenguaje natural (Natural Language Processing, NLP), cubriendo los enfoques clásicos para el procesamiento y análisis de texto, modelos de lenguaje, el modelado de secuencias y la representación de texto, mostrando su aplicación a los problemas habituales de NLP. El curso también cubre una introducción a la aplicación de técnicas de aprendizaje profundo a la NLP y a los fundamentos de los LLMs.

Al final de este curso, los estudiantes podrán:

- Comprender los conceptos y técnicas fundamentales utilizados en NLP.
- Implementar y evaluar varias técnicas de NLP utilizando Python y librerías populares de NLP.
- Aplicar métodos de NLP a problemas del mundo real e interpretar los resultados.

## Resultados de aprendizaje

- SM43 (Analizar texto para descubrir patrones de información relevante.) Analizar texto para descubrir patrones de información relevante.
- SM44 (Desarrollar aplicaciones que requieran un análisis inteligente a nivel semántico del lenguaje natural.) Desarrollar aplicaciones que requieran un análisis inteligente a nivel semántico del lenguaje natural.

## Contenidos

1. Introducción a la lingüística y NLP
2. Procesamiento de texto básico
3. Análisis sintáctico
4. Modelado del lenguaje
5. Representación semántica del texto
6. Pocesamiento de secuencias de texto
7. Aprendizaje profundo para el procesamiento del lenguaje
8. Fundamentos de los LLMs

## Actividades formativas y Metodología

| Título | Horas | ECTS | Resultados de aprendizaje |
|---|---|---|---|
| Trabajo en el proyecto | 50 | 2 | SM44 |
| Clases de teoría | 25 | 1 | SM43, SM44 |
| Sesiones de prácticas | 16 | 0,64 | SM43, SM44 |
| Estudio individual | 24 | 0,96 | SM43 |
| Resolución de problemas (individual) | 25 | 1 | SM43, SM44 |
| Sesiones de proyecto | 4 | 0,16 | SM43, SM44 |

Habrá tres tipos de actividades docentes: clases teóricas, resolución de ejercicios prácticos de forma individual (problemas) y desarrollo de un proyecto en pequeños grupos de 2-3 alumnos.

1. Clases de teoría: Presentación de los contenidos teóricos de la asignatura. Para cada uno de los temas estudiados se exponen los principales conceptos teóricos y formulación matemática, así como las correspondientes soluciones algorítmicas.
2. Sesiones de laboratorio: Las sesiones de laboratorio tienen como objetivo facilitar la interacción y reforzar la comprensión de los temas tratados en las clases de teoría. Durante las sesiones de laboratorio abordaremos dos tipos de actividades: la resolución de ejercicios prácticos (problemas) y el seguimiento y presentación de proyectos.

2.1. Problemas: Se utilizará un conjunto de problemas para trabajar proporcionados en cuadernos Jupyter que ejemplifican los detalles de codificación de los conceptos expuestos durante las clases de teoría. El trabajo de los problemas se iniciará en clase y deberá ser completado en casa. Se requerirá que los estudiantes hagan envíos regulares de su trabajo, que constituirçán a carpeta de problemas (portfolio). En algunas sesiones de clase se podrán realizar actividades evaluativas con respecto a los problemas entregados.

2.2. Proyecto: Se realizará un proyecto durante el semestre, donde los alumnos deberán resolver un problema específico de cierta complejidad. El proyecto se resolverá en pequeños grupos de 2-3 alumnos, donde cada miembro del grupo deberá aportar una parte y juntarla con el resto para obtener la solución final. Estos grupos de trabajo deberán mantenerse hastael final del semestre y deberán autogestionarse en cuanto a distribución de roles, planificación del trabajo, asignación de tareas, gestión de los recursos disponibles, conflictos, etc. Para desarrollar el proyecto, los grupos trabajarán de forma autónoma, mientras que las sesiones prácticas se utilizarán (1) para que el profesor presente el tema del proyecto y discuta posibles enfoques, (2) para el seguimiento del estado del proyecto y (3) para que los equipos presenten sus resultados finales.

Las actividades anteriores se complementarán con un sistema de tutorías y consultas fuera del horario de clases.

Toda la información de la asignatura y los documentos relacionados que necesiten los alumnos estarán

disponibles en el campus virtual (cv.uab.cat).

Nota:se reservarán 15 minutos de una clase dentro del calendario establecido por el centro o por la titulación para que el alumnado rellene las encuestas de evaluación de la actuación del profesorado y de evaluación de la asignatura o módulo.

Nota: se reservarán 15 minutos de una clase dentro del calendario establecido por el centro o por la titulación para que el alumnado rellene las encuestas de evaluación de la actuación del profesorado y de evaluación de la asignatura o módulo.

## Evaluación

### Actividades de evaluación continuada

| Título | Peso | Horas | ECTS | Resultados de aprendizaje |
|---|---|---|---|---|
| Exámenes | 40 | 4 | 0,16 | SM43 |
| Resolución de problemas | 20 | 0 | 0 | SM43, SM44 |
| Proyecto | 40 | 2 | 0,08 | SM44 |

La evaluación se realiza forma continua, por lo que no se ofrece la posibilidad de evaluación única. Para evaluar el nivel de aprendizaje de los alumnos se establece una fórmula que combina la adquisición de conocimientos, la capacidad de resolución de problemas y la capacidad de trabajo en equipo, así como la presentación de los resultados obtenidos.

Nota final

La nota final se calcula de la siguiente forma y en función de las diferentes actividades que se realicen:

Nota final = 0,4 * Nota de teoría + 0,2 * Nota de problemas + 0,4 * Nota de proyecto

Esta fórmula se aplicará siempre que la nota de teoría y del proyecto sea superior a 5. No hay restricción en la nota de los problemas. Si al realizar el cálculo de la fórmula se obtiene >= 5 pero el alumno no alcanza el mínimo exigido en alguna de las actividades de evaluación, entonces se otorgará una calificación final de 4,5.

Nota de teoría

La nota de teoría tiene como objetivo evaluar las capacidades individuales del alumno en cuanto a los contenidos teóricos de la asignatura. Esto se hace de forma continua durante el curso a través de dos exámenes parciales:

Nota de Teoría = 0.5 * Nota Examen 1 + 0.5 * Nota Examen 2

El examen parcial (Examen 1) se realiza a mitad del cuatrimestre y sirve para eliminar parte de la materia si se aprueba. El examen final (Examen 2) se realiza al finalizar el cuatrimestre y sirve para eliminar el resto de asignatura si se aprueba.

Para aprobar la parte de teoría se requerirá que las notas de los exámenes parciales 1 y 2 sean ambas superiores a 4,5 y la media superior a 5,0.

Examen de recuperación: En caso de que la nota de teoría no alcance el nivel adecuado para aprobar, los alumnos podrán realizar un examen de recuperación, destinado a recuperar la parte suspendida (parcial 1, 2 o ambas) del proceso de evaluación continua.

Nota de problemas

El objetivo de los problemas es que el alumno se familiarice con la aplicación práctica de los conceptos teóricos. La entrega regular de los problemas resueltos se utilizará como evidencia de este trabajo.

Para obtener la calificación de los problemas es necesario que más del 50% de los mismos se entreguen durante el semestre. En caso contrario, la nota de los problemas será 0.

Habrá una evaluación individual del trabajo del estudiante en los problemas que puede consistir en preguntas en examen sobre los problemas u otros tipos de actividades evaluativas durante las sesiones de clase. La nota final de los problemas será la combinación del dossier de problemas y de la evaluación individual.

Nota de problemas = 0.5 * Evaluación del portfolio + 0.5 * Evaluación individual

Nota del proyecto

El proyecto tiene un peso esencial en la nota global de la asignatura. El desarrollo del proyecto requiere que los alumnos trabajen en grupo y diseñen una solución integral al reto definido. Además, los estudiantes deben demostrar sus habilidades de trabajo en equipo y presentar los resultados en clase.

El proyecto se evalúa a través de su entregable, una presentación oral que los estudiantes realizarán en clase y un proceso de evaluación individual. La participación de los estudiantes en las tres actividades (preparación del entregable, presentación y evaluación individual) es necesaria para obtener la calificación de los proyectos. La calificación se calcula dela siguiente manera:

Nota del Proyecto = 0.6 * Evaluación Entregables + 0.2 * Evaluación Presentación + 0.2 *Evaluación individual

Si al realizar el cálculo anterior se obtiene >= 5 pero el estudiante no participó en alguna de las actividades (entregables, presentación, evaluación individual), entonces se otorgará una calificación final de 4.5 al proyecto correspondiente.

En caso de que se presente el entregable, pero la nota final del proyecto no alcance el mínimo de 5, habrá recuperación del proyecto. En caso de no presentar el entregable o considerarlo copiado, no habrá recuperación y la asignatura será considerada suspendida. La nota máxima del proyecto que se puede obtener en caso de recuperación es de 7.

Los alumnos repetidores pueden convalidar la nota del proyecto del curso anterior. En este caso, independientemente de la nota obtenida en el curso anterior la nota del proyecto será un 5.

Notas importantes

Sin perjuicio de otras medidas disciplinarias que se estimen oportunas, y de conformidad con la normativa académica vigente, las actividades de evaluación se suspenderán con cero (0) cuando un alumno incurra en irregularidades académicas que puedan alterar dicha evaluación (por ejemplo, plagiar, copiar, dejar copiar, ...). Las actividades de evaluación calificadas de esta forma y por este procedimiento no serán recuperables. Si necesitas superar alguna de estas actividades de evaluación para aprobar la asignatura, se suspenderá directamente esta asignatura, sin posibilidad de recuperarla en el mismo curso.

Para esta asignatura se permite el uso de tecnologías de Inteligencia Artificial generativa de forma controlada. El estudiante deberá identificar claramente qué partes de su trabajo han sido realizadas con soporte de herramientas de IA generativa y, en todo caso, deberá ser capaz de entender, explicar y justificar el trabajo realizado. La no transparencia del uso de la IA generativa se considerará falta de honestidad académica y puede comportar una penalización parcial o total en la nota de la actividad.

En caso de que el estudiante no entregue la solución de problemas, no asista a las sesiones de presentación de proyectos durante las sesiones de laboratorio y no realice ningún examen, la calificación correspondiente será \"no evaluable\". En otro caso, los \"no shows\" cuentan como 0 para el cálculo del promedio ponderado.

Para aprobar la asignatura con matrícula de honor, la nota final obtenida deberá ser igual o superior a 9 puntos. Debido a que el número de alumnos con esta distinción no puede exceder el 5% del total de alumnos matriculados en el curso, se otorga a quien obtenga la nota final más alta. En caso de empate, se tendrán en

cuenta los resultados de los exámenes parciales.

## Bibliografía


- D. Jurafsky, JH Martín. *Procesamiento del habla y el lenguaje* . Tercera edicion. 2021 < https://web.stanford.edu/~jurafsky/slp3/ >
- J. Eisenstein. *Procesamiento del lenguaje natural* . 2018. Prensa del MIT
- H. Lane, C. Howard, HM Hapke. *Procesamiento del lenguaje natural en acción* . 2019. Publicaciones de Manning
- Kenny, Dorothy, ed. *Traducción automática para todos* . Prensa científica lingüística, 2022. < https://langsci-press.org/catalog/book/342 > _ _
- Rowe, Bruce M. y Diane P. Levine. *Una breve introducción a la lingüística* . Routledge, 2018.


## Software

Para los problemas y proyectos del curso utilizaremos Python, junto con algunas librerías de Python para NLP que se especificarán durante el curso.

### Grupos e idiomas de la asignatura

La información proporcionada es provisional hasta el 30 de noviembre. A partir de esta fecha, podrá consultar el idioma de cada grupo a través de este enlace. Para acceder a la información, será necesario introducir el CÓDIGO de la asignatura