



## DATOS IDENTIFICATIVOS

### Comprensión del lenguaje natural

|                     |  |            |       |              |
|---------------------|--|------------|-------|--------------|
| Asignatura          | Comprensión del lenguaje natural   |            |       |              |
| Código              | O06M193V01104  |            |       |              |
| Titulación          | Máster universitario en Inteligencia artificial  |            |       |              |
| Descriptores        | Creditos ECTS  | Seleccione | Curso | Cuatrimestre |
|                     | 6  | OB         | 1     | 1c           |
| Lengua Impartición  | Inglés   |            |       |              |
| Departamento        |  |            |       |              |
| Coordinador/a       | Darriba Bilbao, Víctor Manuel  |            |       |              |
| Profesorado         | Darriba Bilbao, Víctor Manuel  |            |       |              |
| Correo-e            | darriba@uvigo.es   |            |       |              |
| Web                 | <a href="http://guiadocente.udc.es/guia_docent/index.php?centre=614&amp;ensenyament=614544&amp;consulta=assignatures&amp;ny_academic=2023_24">http://guiadocente.udc.es/guia_docent/index.php?centre=614&amp;ensenyament=614544&amp;consulta=assignatures&amp;ny_academic=2023_24</a>  |            |       |              |
| Descripción general | La asignatura introduce los conceptos y técnicas básicas asociadas al procesamiento del lenguaje natural, punto de partida para el diseño de entornos de explotación de información y de diálogo basados en el lenguaje humano, tanto a nivel léxico como sintáctico, semántico y pragmático. El objetivo es introducir al estudiante en la complejidad inherente al análisis del lenguaje natural humano, fundamentalmente asociada a la ambigüedad y dependencias contextuales que presenta, y en el diseño de estructuras de datos y algoritmos que permitan su tratamiento práctico. |            |       |              |

## Resultados de Formación y Aprendizaje

|        |   |
|--------|---|
| Código |   |
| A1     | CB6 - Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación   |
| A2     | CB7 - Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio |
| A5     | CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.  |
| B1     | Mantener y extender planteamientos teóricos fundados para permitir la introducción y explotación de tecnologías nuevas y avanzadas en el campo de la Inteligencia Artificial.   |
| B3     | Buscar y seleccionar la información útil necesaria para resolver problemas complejos, manejando con soltura las fuentes bibliográficas del campo.   |
| B4     | Elaborar adecuadamente y con cierta originalidad composiciones escritas o argumentos motivados, redactar planes, proyectos de trabajo, artículos científicos y formular hipótesis razonables en el campo.                                     |
| C1     | Comprensión y dominio de técnicas para el procesado de textos en lenguaje natural.  |
| C2     | Comprensión y dominio de los fundamentos y técnicas de procesamiento semántico de documentos enlazados, estructurados y no estructurados, y de la representación de su contenido.   |
| C3     | Comprensión y conocimiento de las técnicas de representación y procesado de conocimiento mediante ontologías, grafos y RDF, así como de las herramientas asociadas a las mismas.  |
| D2     | Dominar la expresión y la comprensión de forma oral y escrita de un idioma extranjero.  |
| D3     | Utilizar las herramientas básicas de las tecnologías de la información y las comunicaciones (TIC) necesarias para el ejercicio de su profesión y para el aprendizaje a lo largo de su vida.   |
| D7     | Desarrollar la capacidad de trabajar en equipos interdisciplinares o transdisciplinares, para ofrecer propuestas que contribuyan a un desarrollo sostenible ambiental, económico, político y social.  |
| D8     | Valorar la importancia que tiene la investigación, la innovación y el desarrollo tecnológico en el avance socioeconómico y cultural de la sociedad  |

## Resultados previstos en la materia

| Resultados previstos en la materia   | Resultados de Formación y Aprendizaje                                      |
|--|--|
| Conocer, comprender y analizar la representación formal de diversos fenómenos léxicos, sintácticos y semánticos del lenguaje natural.                        | A1<br>A5<br>B1<br>B3<br>B4<br>C1<br>D2<br>D8                               |
| Conocer, comprender y saber usar las tecnologías, marcos y librerías para la construcción de sistemas de procesamiento del lenguaje natural.                 | A1<br>A2<br>A5<br>B3<br>B4<br>C1<br>C2<br>D2<br>D3<br>D7                   |
| Diseñar, implementar y saber usar algoritmos y estructuras de datos para tratar y dar soporte a los diversos fenómenos característicos del lenguaje natural. | A1<br>A2<br>A5<br>B1<br>B3<br>B4<br>C1<br>C2<br>C3<br>D2<br>D3<br>D7<br>D8 |
| Conocer, comprender y analizar las técnicas de procesamiento del lenguaje natural para el procesado y desambiguación a nivel léxico, sintáctico y semántico. | A1<br>A2<br>A5<br>B1<br>B3<br>B4<br>C1<br>C2<br>C3<br>D2<br>D3<br>D7<br>D8 |
| Conocer y comprender los problemas que plantea la ambigüedad e imprecisión en las fuentes de datos en lenguaje natural y técnicas para resolverlos.          | A1<br>A2<br>A5<br>B1<br>B3<br>B4<br>C1<br>C3<br>D2<br>D3<br>D7<br>D8       |

## Contenidos

### Tema

Introducción.

Niveles de análisis.

Ambigüedad y dependencias contextuales.

|                      |   |
|----------------------|---|
| Análisis léxico.     | Segmentación.<br>Diccionarios y tesauros.<br>Técnicas de etiquetación morfosintáctica.  |
| Análisis sintáctico. | Gramáticas algebraicas.<br>Gramáticas suavemente sensibles al contexto.<br>Gramáticas de dependencias.<br>Gramáticas probabilísticas. |
| Análisis semántico.  | Semántica léxica.<br>Dependencias semánticas.<br>Grafos semánticos.   |

### Planificación

|                               | Horas en clase | Horas fuera de clase | Horas totales |
|-------------------------------|----------------|----------------------|---------------|
| Lección magistral             | 21             | 21                   | 42            |
| Prácticas de laboratorio      | 14             | 48                   | 62            |
| Resolución de problemas       | 9              | 25                   | 34            |
| Examen de preguntas objetivas | 3              | 9                    | 12            |

\*Los datos que aparecen en la tabla de planificación son de carácter orientativo, considerando la heterogeneidad de alumnado

### Metodologías

|                          | Descripción  |
|--------------------------|--|
| Lección magistral        | Clases teóricas, en las que se expone el contenido de cada tema. El alumno dispondrá de copias de las transparencias con anterioridad y el profesor promoverá una actitud activa, realizando preguntas que permitan aclarar aspectos concretos y dejando cuestiones abiertas para la reflexión del alumno. |
| Prácticas de laboratorio | Clases prácticas con uso de ordenador, que permiten al alumno familiarizarse desde un punto de vista práctico con las cuestiones expuestas en las clases teóricas.   |
| Resolución de problemas  | Aprendizaje basado en problemas, seminarios, estudio de casos y proyectos.   |

### Atención personalizada

| Metodologías             | Descripción  |
|--------------------------|--|
| Lección magistral        | El profesorado atenderá al alumnado en sesiones de tutorías individualizadas, dedicadas a la orientación en el estudio y a la resolución de dudas sobre los contenidos, trabajos y actividades de la asignatura. |
| Prácticas de laboratorio | El profesorado atenderá al alumnado en sesiones de tutorías individualizadas, dedicadas a la orientación en el estudio y a la resolución de dudas sobre los contenidos, trabajos y actividades de la asignatura. |
| Resolución de problemas  | El profesorado atenderá al alumnado en sesiones de tutorías individualizadas, dedicadas a la orientación en el estudio y a la resolución de dudas sobre los contenidos, trabajos y actividades de la asignatura. |

### Evaluación

|                               | Descripción  | Calificación | Resultados de Formación y Aprendizaje        |
|-------------------------------|--|--------------|--|
| Prácticas de laboratorio      | Las entregas de las prácticas deben realizarse dentro del plazo establecido en el campus virtual y deben seguir las especificaciones indicadas en el enunciado tanto para su presentación como su defensa. | 50           | A1 B3 C1 D2<br>A2 B4 C2 D3<br>A5 C3 D7<br>D8 |
| Examen de preguntas objetivas | Realización obligatoria. Se evaluará el dominio de los conocimientos teóricos y operativos de la asignatura.   | 50           | A1 B1 C1 D2<br>A2 C2<br>C3                   |

### Otros comentarios sobre la Evaluación

## **CRITERIOS DE EVALUACIÓN PARA TODOS LOS ESTUDIANTES EN TODAS LAS CONVOCATORIAS**

Los alumnos deberán alcanzar al menos un 40% de la máxima nota cada parte (teoría, práctica) y en cualquier caso la suma de ambas partes debe superar un 5 para superar la asignatura. Si no se cumple alguno de los requisitos anteriores, la nota de la convocatoria se establecerá de acuerdo a la menor nota obtenida.

En caso de no alcanzar el mínimo en alguna de las partes, el alumno tendrá una segunda oportunidad en la que solamente se le exigirá la entrega de dicha parte.

Las entregas de las prácticas deben realizarse dentro del plazo establecido en el campus virtual y deben seguir las especificaciones indicadas en el enunciado tanto para su presentación como su defensa.

Tendrá la condición de "Presentado" quien entregue todas las prácticas obligatorias o concurra a la prueba objetiva en el período oficial de evaluación.

En el caso de realización fraudulenta de ejercicios o pruebas, se aplicará la Normativa de evaluación del rendimiento académico de los estudiantes y de revisión de calificaciones. En aplicación de la normativa correspondiente sobre plagio, la copia total o parcial de algún ejercicio de prácticas o de teoría supondrá el suspenso en las dos oportunidades del curso, con la calificación de 0,0 en ambos casos.

## **FECHAS DE EVALUACIÓN**

Las fechas oficiales de examen de las diferentes convocatorias, se encuentran publicadas en la página web de la ESEI: <https://esei.uvigo.es/docencia/exames/>

## **CONSULTA/SOLICITUD DE TUTORÍAS**

Las tutorías pueden consultarse a través de la página personal del profesorado, accesible a través de <https://esei.uvigo.es/docencia/profesorado/>

---

### **Fuentes de información**

#### **Bibliografía Básica**

Manning, C., & Schütze, H., **Foundations of statistical natural language processing**, 978-0262133609, 1, MIT Press, 1999

Goldberg, Y., **Neural network methods for natural language processing. Synthesis lectures on human language technologies**, 978-1627052986, 1, Morgan Claypool, 2017

Eisenstein, J., **Introduction to Natural Language Processing**, 978-0262042840, 1, MIT Press, 2019

Jurafsky, D. & Martin, J. H., **Speech and Language Processing**, 978-0131873216, 3 (draft), <https://web.stanford.edu/~jurafsky/slp3/>, 2022

Jurafsky, D. & Martin, J. H., **Speech and Language Processing**, 978-0131873216, 2, Prentice Hall, 2008

Indurkha, N. & Damerau, F.J. (Eds.), **Handbook of Natural Language Processing**, 978-1420085921, 2, Routledge, 2010

#### **Bibliografía Complementaria**

Chollet, F., **Keras: The python deep learning library**, Astrophysics Source Code Library, 2018

Russell, S., Norvig, P., **Artificial Intelligence: A Modern Approach**, 978-0134610993, 4, Pearson, 2022

Manning, C.D., Raghavan, P., Schütze, H., **Introduction to Information Retrieval**, 978-0521865715, 1, Cambridge University Press, 2008

Kübler, S., McDonald, R., & Nivre, J., **Dependency Parsing. Synthesis lectures on human language technologies**, 978-1598295962, 1, Morgan Claypool, 2009

---

### **Recomendaciones**

#### **Asignaturas que continúan el temario**

Inteligencia web y tecnologías semánticas/O06M193V01205

Modelado del lenguaje/O06M193V01204

Minería de textos/O06M193V01302

---

#### **Asignaturas que se recomienda cursar simultáneamente**

Aprendizaje automático I/O06M193V01105

---

### **Otros comentarios**

Se recuerda a todo el alumnado la prohibición del uso de dispositivos móviles en ejercicios, prácticas y exámenes, en cumplimiento del artículo 13.2.d) del Estatuto del Estudiante Universitario, relativo a los deberes del estudiantado universitario, que establece el deber de "Abstenerse de la utilización o cooperación en procedimientos fraudulentos en las pruebas de evaluación, en los trabajos que se realicen o en documentos oficiales de la universidad."