

17-18

MÁSTER UNIVERSITARIO EN I.A.  
AVANZADA: FUNDAMENTOS, MÉTODOS  
Y APLICACIONES

# GUÍA DE ESTUDIO PÚBLICA



## PROCESAMIENTO DEL LENGUAJE NATURAL

CÓDIGO 31101269



Ámbito: GUJ - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección <https://sede.uned.es/valida/>



AG761B5F80BC040F69200CADAFA80BCDE

# ÍNDICE

|                                                            |  |
|------------------------------------------------------------|--|
| PRESENTACIÓN Y CONTEXTUALIZACIÓN                           |  |
| REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA |  |
| EQUIPO DOCENTE                                             |  |
| HORARIO DE ATENCIÓN AL ESTUDIANTE                          |  |
| COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE                    |  |
| RESULTADOS DE APRENDIZAJE                                  |  |
| CONTENIDOS                                                 |  |
| METODOLOGÍA                                                |  |
| SISTEMA DE EVALUACIÓN                                      |  |
| BIBLIOGRAFÍA BÁSICA                                        |  |
| BIBLIOGRAFÍA COMPLEMENTARIA                                |  |
| RECURSOS DE APOYO Y WEBGRAFÍA                              |  |



Nombre de la asignatura  
Código  
Curso académico  
Títulos en que se imparte

PROCESAMIENTO DEL LENGUAJE NATURAL  
31101269  
2017/2018  
MÁSTER UNIVERSITARIO EN I.A. AVANZADA:  
FUNDAMENTOS, MÉTODOS Y APLICACIONES (máster seleccionado) /  
MÁSTER UNIVERSITARIO EN LENGUAJES Y SISTEMAS  
INFORMÁTICOS  
CONTENIDOS  
6  
150.0  
ANUAL  
CASTELLANO

## PRESENTACIÓN Y CONTEXTUALIZACIÓN

Este curso introductorio al procesamiento computacional del lenguaje natural aborda el diseño y la construcción de programas que pueden tratar, comprender y generar lenguaje natural. Se estudiarán los problemas y soluciones (modelos y técnicas) básicas en los niveles sintáctico, semántico y pragmático. Un capítulo de introducción y otro de áreas de aplicación, situarán la asignatura desde una perspectiva histórica, y permitirán conocer el estado actual de las realizaciones en este campo. La realización de un proyecto permitirá poner en práctica los conocimientos adquiridos.

Ficha técnica:

|                          |          |
|--------------------------|----------|
| Tipo                     | Optativa |
|                          |          |
| Créditos/horas totales   | 6/150    |
| Horas de estudio teórico | 45       |
| Horas de prácticas       | 45       |
| Proyecto                 | 60       |

## REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA

Es importante una lectura fluida del inglés y disponer de conexión a internet. En cuanto a contenidos, esta asignatura tiene relación estrecha con las asignaturas de la carrera de Ingeniería Informática relativas a Teoría de autómatas, Procesadores de Lenguaje, e Introducción a la Inteligencia Artificial, que proporcionan la base en cuanto a formalismos y técnicas computacionales. Así mismo las asignaturas de programación, y especialmente aquellas en que se estudian paradigmas declarativos, constituyen un complemento



interesante para cursar Procesamiento de Lenguaje Natural. El lenguaje de programación que se utilizará es Python.

## EQUIPO DOCENTE

Nombre y Apellidos  
Correo Electrónico  
Teléfono  
Facultad  
Departamento

MARIA FELISA VERDEJO MAILLO  
felisa@lsi.uned.es  
91398-6484  
ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA  
LENGUAJES Y SISTEMAS INFORMÁTICOS

Nombre y Apellidos  
Correo Electrónico  
Teléfono  
Facultad  
Departamento

ENRIQUE AMIGO CABRERA  
enrique@lsi.uned.es  
91398-8651  
ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA  
LENGUAJES Y SISTEMAS INFORMÁTICOS

## HORARIO DE ATENCIÓN AL ESTUDIANTE

La tutorización de los alumnos se llevará a cabo a través de una plataforma de aprendizaje en internet, que incluye comunicación electrónica con el equipo docente, y áreas de compartición de trabajo. Los trabajos que se proponen se evalúan a lo largo del curso.

## COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE

## RESULTADOS DE APRENDIZAJE

En la primera parte del curso, mediante el estudio de la bibliografía el alumno adquirirá una visión amplia de las técnicas de procesamiento de lenguaje natural en los niveles léxico, sintáctico y semántico y sus aplicaciones. Los conocimientos adquiridos a nivel teórico se pondrán en práctica mediante la realización de ejercicios con una herramienta, NLTK, de libre disposición que los alumnos utilizarán para la elaboración de un analizador morfológico y un analizador sintáctico y semántico sobre un subdominio abordable del lenguaje. Paralelamente, los conocimientos adquiridos a nivel global y la capacidad de síntesis se pondrán en práctica mediante el desarrollo de una serie de resúmenes (el primero de ellos guiado por preguntas). En la segunda parte del curso el alumno adquirirá la destreza necesaria para elaborar un sistema de procesamiento de lenguaje orientado a una tarea específica. Con este curso el alumno asimilará tanto el potencial de las técnicas existentes de procesamiento de lenguaje como de sus limitaciones, siendo capaz de analizar en qué casos es factible aplicar estas técnicas en la resolución de un problema.



## CONTENIDOS

## METODOLOGÍA

El curso consta de un conjunto de temas cuyo estudio se realiza con la siguiente metodología dentro de un paradigma de construcción de conocimiento:

Para cada tema, el alumno debe acceder al material propuesto por el equipo docente. Este material consta de:

- Bibliografía básica común a todos los temas. Se trata un libro con un conocimiento ya estructurado facilitando la introducción del alumno en la materia.
- Artículos científicos. Se propone la lectura de algunos artículos de carácter científico. Su contenido es más específico. Aparte de conocer su contenido, el alumno se familiarizará con la estructura y formato que deben seguir los textos de estas características.
- Enlaces web que apuntan a recursos y herramientas relacionados con el tema.

A partir de este material (y con la guía de unas preguntas para el primero), el alumno debe realizar tres síntesis (de 10-30 páginas), en forma de artículos, correspondiendo cada una a un bloque de temas, con el objetivo de sintetizar el conocimiento que ha adquirido. La elaboración del artículo se dirige a:

- Estimular la lectura detenida del material propuesto.
- Provocar la necesidad de buscar información que complete el material propuesto inicialmente. Esta búsqueda es un entrenamiento necesario en la formación del alumno como investigador. Con cada trabajo tendrá mayor capacidad para encontrar y discriminar fuentes de información relevantes, requisito para desarrollar cualquier trabajo de investigación posterior.
- Estimular una reflexión sobre el material estudiado, necesaria para poder realizar una síntesis de calidad.
- Adquirir destreza para la escritura de documentos técnicos.

Junto con la elaboración de artículos de síntesis, el alumno deberá realizar una serie de ejercicios para los que dispondrá de la herramienta NLTK. De esta forma, el alumno podrá centrarse en aspectos del lenguaje y diseño de gramáticas sin necesidad de implementar desde un principio autómatas y mecanismos de unificación. Para manejar esta herramienta es conveniente tener conocimientos del lenguaje de programación Python, por lo que la primera semana del curso se dedicará a este tema.

Los últimos meses del curso se dirigen a afianzar los conocimientos adquiridos mediante la elaboración de un proyecto en el que se pondrán en práctica las técnicas aprendidas con



datos y herramientas similares a los utilizados en la practica profesional

## SISTEMA DE EVALUACIÓN

## BIBLIOGRAFÍA BÁSICA

ISBN(13):9780135041963

Título:SPEECH AND LANGUAGE PROCESSING (segunda)

Autor/es:Jurafsky, Daniel ; Martin, James H. ;

Editorial:PEARSON EDUCATION

El libro base (segunda edición) es imprescindible. La primera edición del libro de bibliografía básica no incluye algunos contenidos del curso.

Para los ejercicios prácticos:

Libro: Natural language processsing with Python, disponible en

<http://victoria.lviv.ua/html/fl5/NaturalLanguageProcessingWithPython.pdf>

### Herramientas

Python (lenguaje de programación)

<https://docs.python.org/3/>

NLTK 3.0

<http://www.nltk.org/>

## BIBLIOGRAFÍA COMPLEMENTARIA

En el entorno virtual de la asignatura se pone a disposición de los alumnos material de estudio complementario (artículos, recopilaciones y referencias a otro material disponible en la web)

## RECURSOS DE APOYO Y WEBGRAFÍA

La plataforma de aprendizaje en internet, proporcionará interfaz de interacción entre el alumno y sus profesores. Esta plataforma de e-Learning y colaboración permite impartir y recibir formación, gestionar y compartir documentos, crear y participar en comunidades temáticas, así como realizar proyectos online.

Además, los alumnos tendrán que descargarse la herramienta NLTK, para la realización de ejercicios prácticos

¿Qué es NLTK? Una presentación breve

<http://desilinguist.org/pdf/crossroads.pdf>

Para el proyecto se facilitará el acceso a los datos y herramientas necesarias para llevarlo a cabo.



---

## IGUALDAD DE GÉNERO

En coherencia con el valor asumido de la igualdad de género, todas las denominaciones que en esta Guía hacen referencia a órganos de gobierno unipersonales, de representación, o miembros de la comunidad universitaria y se efectúan en género masculino, cuando no hayan sido sustituido por términos genéricos, se entenderán hechas indistintamente en género femenino o masculino, según el sexo del titular que los desempeñe.

