



Guía Docente				
Datos Identificativos				2012/13
Asignatura (*)	Procesamento Avanzado da Linguaxe Natural		Código	614434011
Titulación	Mestrado Universitario en Computación			
Descritores				
Ciclo	Período	Curso	Tipo	Créditos
Mestrado Oficial	2º cuatrimestre	Primeiro	Optativa	3
Idioma	Castelán			
Prerrequisitos				
Departamento	Computación			
Coordinación	Alonso Pardo, Miguel angel		Correo electrónico	miguel.alonso@udc.es
Profesorado	Alonso Pardo, Miguel angel Gómez Rodríguez, Carlos		Correo electrónico	miguel.alonso@udc.es carlos.gomez@udc.es
Web	www.dc.fi.udc.es/muc/?q=es/courses/PALN			
Descrición xeral	Esta asignatura aborda el Procesamiento del Lenguaje Natural (NLP, Natural Language Processing), la rama de las ciencias computacionales encargada del diseño e implementación de los elementos software necesarios para el tratamiento computacional del lenguaje natural, entendiendo como tal toda lengua je humano, en contraposición a los lenguajes formales propios del ámbito lógico, matemático, o computacional. El objetivo último que se persigue, y que todavía está muy lejos de conseguirse, es el de la comprensión del lenguaje humano por parte de la computadora.			

Competencias da titulación	
Código	Competencias da titulación
A1	Adquirir coñecementos de Lóxicas Computacionais e as súas principais aplicacións a outras áreas específicas de investigación en Computación tales como Raonamento Automático, Representación do Coñemento, Razoamento Temporal e Espacial, Sistemas Multiaxente, Web semántica, Verificación Formal, etc.
A2	Comprender os conceptos básicos da aprendizaxe computacional, as diferentes técnicas dispoñibles e o seu ámbito de aplicabilidade. Ser capaz de aplicar as distintas técnicas de aprendizaxe empregando unha metodoloxía axeitada.
A4	Posuír unha ampla comprensión dos sistemas de Xestión da Información, desde os aspectos máis técnicos como as Estructuras de Datos Compactas e os correspondentes algoritmos de uso, ata as máis avanzadas técnicas de Recuperación da Información, Extracción de Información e Procura de Respostas.
B1	Ser capaz de formular xuízos a partir dunha información que, sendo incompleta ou limitada, inclúa reflexións sobre as responsabilidades sociais e éticas vinculadas á aplicación dos seus coñecementos e xuízos.
B2	Destreza na adquisición do coñecemento, análise do estado da arte e bibliografía relevante nunha área de investigación.
B3	Capacidade para identificar problemas e formular adecuadamente as hipóteses a contrastar seguindo unha metodoloxía científica.
B4	Aplicación do método científico mediante análise empírico das hipóteses formuladas ou mediante demostración formal, no caso de propiedades matemáticas. Destreza no deseño de experimentos e a análise de resultados.
B5	Aptitude para a correcta elaboración e redacción de publicacións científicas tales como artigos de revista ou informes técnicos.
B7	Acostumarse ó uso do inglés como principal idioma de adquisición e transmisión de coñecemento científico e de investigación.
C1	Expresarse correctamente, tanto de forma oral coma escrita, nas linguas oficiais da comunidade autónoma.
C2	Dominar a expresión e a comprensión de forma oral e escrita dun idioma estranxeiro.
C3	Utilizar as ferramentas básicas das tecnoloxías da información e as comunicacións (TIC) necesarias para o exercicio da súa profesión e para a aprendizaxe ao longo da súa vida.
C4	Desenvolverse para o exercicio dunha cidadanía aberta, culta, crítica, comprometida, democrática e solidaria, capaz de analizar a realidade, diagnosticar problemas, formular e implantar solucións baseadas no coñecemento e orientadas ao ben común.
C5	Entender a importancia da cultura emprendedora e coñecer os medios ao alcance das persoas emprendedoras.
C6	Valorar criticamente o coñecemento, a tecnoloxía e a información dispoñible para resolver os problemas cos que deben enfrontarse.
C7	Asumir como profesional e cidadán a importancia da aprendizaxe ao longo da vida.
C8	Valorar a importancia que ten a investigación, a innovación e o desenvolvemento tecnolóxico no avance socioeconómico e cultural da sociedade.



## Resultados da aprendizaxe

Resultados da aprendizaxe		Competencias de titulación	
Competencias de materia (Resultados de aprendizaxe)		Competencias de titulación	
Comprender la descripción lingüística de los fenómenos presentes en el lenguaje natural		BI3	
Reconocer los fenómenos del lenguaje que son tratables actualmente y aquellos que no lo son.		BI2 BI3	CM8
Conocer los algoritmos, técnicas y métodos más actuales utilizados en el procesamiento del lenguaje natural.	AI1 AI2 AI4		CM3 CM6 CM7 CM8
Diseñar e implementar algoritmos y estructuras de datos para el tratamiento de diversos fenómenos de los lenguajes humanos.	AI2 AI4	BI3 BI4	CM3 CM6 CM8
Usar las técnicas y métodos del procesamiento del lenguaje natural para resolver problemas reales de comunicación hombre-máquina.		BI1 BI2 BI3 BI4	CM3 CM5 CM6 CM8
Asumir la complejidad del lenguaje humano y las limitaciones de los ordenadores para tratarlo.			CM4 CM5 CM6 CM7 CM8
Aceptar soluciones parciales a un problema ante la imposibilidad de obtener soluciones completas.		BI3	CM3 CM6 CM8
Rechazar la compartimentación del saber en áreas cerradas (por ejemplo, ciencias contraposición a humanidades)			CM4 CM8
Presentar los resultados obtenidos conforme a las normas establecidas en el campo del Procesamiento del Lenguaje Natural		BI1 BI5 BI7	CM1 CM2

## Contidos

Temas	Subtemas
Introdución al procesamiento avanzado del lenguaje natural	Recursos lingüísticos Preprocesamiento
Análisis a nivel de palabra	Etiquetación morfosintáctica - métodos basados en reglas - métodos estadísticos (modelos de markov ocultos, máxima entropía) Desambigüación del sentido de las palabras
Análisis sintáctico basado en constituyentes	Esquemas de análisis sintáctico Análisis con formalismos gramaticales independientes del contexto Análisis con formalismos gramaticales suavemente sensibles al contexto Análisis sintáctico con reparación automática de errores
Análisis sintáctico basado en dependencias	Análisis de dependencias basados en grafos Análisis de dependencias basado en transiciones Complejidad computacional y relaciones con otros modelos y formalismos
Análisis semántico	Formalismos de representación semántica; Correferencia y anáfora; Análisis semántico composicional



Traducción automática	Modelos de traducción automática Corpus paralelos y comparables
-----------------------	--

Planificación			
Metodoloxías / probas	Horas presenciais	Horas non presenciais / traballo autónomo	Horas totais
Sesión maxistral	10	10	20
Presentación oral	2	24	26
Traballos tutelados	1	18	19
Atención personalizada	10	0	10

\*Os datos que aparecen na táboa de planificación son de carácter orientativo, considerando a heteroxeneidade do alumnado

Metodoloxías	
Metodoloxías	Descrición
Sesión maxistral	En las clases presenciais de teoría, el profesor realizará una breve descripción de los contenidos temáticos y de los objetivos básicos perseguidos, con el fin de dotar al alumno de una visión global de la materia. Además tratará de establecer interrelaciones con otros conceptos previamente adquiridos, de forma que se pueda establecer una línea temporal, y expondrá la bibliografía recomendada. Seguidamente pasará a desarrollar los contenidos teóricos, utilizando como método la clase magistral.
Presentación oral	Se propondrá al alumno el estudio de un tema concreto. Para ello se lo proporcionará un conjunto de artículos, cuyo contenido el alumno deberá ser capaz de asimilar, sintetizar y concretar en una presentación que estará sujeta a las críticas tanto del profesor como de otros alumnos.
Traballos tutelados	Se pretende que el alumno encuentre soluciones para un problema dado. Estas soluciones pueden venir tanto en la forma de la definición de estructuras teóricas como en la implementación de algoritmos. Se impondrán unos puntos de control para fomentar el estudio continuo. El enunciado, que se proporcionará con la suficiente antelación, detallará el problema y las especificaciones, que el alumno deberá respetar estrictamente. Este enunciado se analizará con los alumnos. Posteriormente, la labor del profesor será la de supervisar la actuación del alumno, solucionando dudas, corrigiendo errores de interpretación, sugiriendo lecturas, etc.

Atención personalizada	
Metodoloxías	Descrición
Presentación oral Traballos tutelados	La labor del profesor será, tanto en las clases teóricas como prácticas, la de supervisar el trabajo del alumno, solucionando dudas, corrigiendo errores de interpretación, sugiriendo lecturas, etc., no sólo como grupo, sino también como individuo.

Avaliación		
Metodoloxías	Descrición	Cualificación
Sesión maxistral	Se valorará la participación activa en las clases	20
Presentación oral	Se valorará la capacidad de comprensión del tema propuesto, la capacidad de síntesis y las habilidades de expresión, tanto escritas (realización de la presentación) como orales.	40
Traballos tutelados	Se valorará la profundidad del estudio realizada por el alumno, así como el grado de corrección y completud de la solución propuesta	40

Observacións avaliación



## Fontes de información

<b>Bibliografía básica</b>	<ul style="list-style-type: none"><li>- Christopher D. Manning and Hinrich Schütze (1999.). Foundations of Statistical Natural Language Processing. The MIT Press, Cambridge (Massachusetts) and London (England)</li><li>- Nitin Indurkha and Fred J. Damerau (2010). Handbook of Natural Language Processing (second edition). Chapman Hall/CRC, Boca Raton, FL</li><li>- Carlos Gómez-Rodríguez (2010). Parsing Schemata for Practical Text Analysis. Imperial College Press, London</li><li>- Daniel Jurafsky and James H. Martin (2009). Speech and Language Processing. Pearson Education, Upper Saddle River (New Jersey)</li><li>- Ruslan Mitkov (editor) (2005.). The Oxford Handbook of Computational Linguistics. Oxford University Press, Oxford</li></ul>
<b>Bibliografía complementaria</b>	

## Recomendacións

### Materias que se recomenda ter cursado previamente

Extracción e Recuperación da Información/614434003

### Materias que se recomenda cursar simultaneamente

Programación Lóxica e Representación do Coñecemento/614434012

### Materias que continúan o temario

### Observacións

(\*A Guía docente é o documento onde se visualiza a proposta académica da UDC. Este documento é público e non se pode modificar, salvo casos excepcionais baixo a revisión do órgano competente dacordo coa normativa vixente que establece o proceso de elaboración de guías