

1. Datos descriptivos de la asignatura

Nombre de la asignatura: Aplicaciones Inteligentes para la Web

Curso Académico: 2011/12 **Curso:** Tercer i Quart

Trimestre: Primer

Titulació: Grau en Enginyeria en Informàtica; Grau en Enginyeria de Sistemes Audiovisuals

Código asignatura: 21449 i 21635

Número de créditos ECTS: 4

Horas dedicación estudiante: 100

Lengua de docència: Castellà

Profesorado: Nadjat Bouayad Agha i Horacio Saggion

2. Presentación de la asignatura

En esta asignatura se estudiarán técnicas y algoritmos para el desarrollo de aplicaciones inteligentes para la Web. Se estudiarán algoritmos tanto supervisados como no supervisados para el tratamiento de datos de la Web y se analizarán las ventajas y desventajas de los métodos presentados. Asimismo se estudiarán aplicaciones como la extracción de información, la minería de textos y opiniones, y el resumen automático. También se incluirán tópicos avanzados como la Web Semántica. Se requiere conocimiento de algoritmos y estructuras de datos y lógica para computación. Se requieren conocimientos de programación Java para el desarrollo de las prácticas.

3. Competencias a assolir a l'assignatura

Competencias transversales	Competencias específicas
<i>Instrumentales</i> 1. Análisis y síntesis 2. Organización y planificación 3. Búsqueda y gestión de información 4. Capacidad de comunicación.	 1. Fundamentos de bases de datos no estructuradas y técnicas relacionadas con clasificación y recuperación de información. 2. Conocimiento del funcionamiento de redes, en particular la Internet. 3. Técnicas de búsqueda, planificación, y razonamiento inteligente.
<i>Interpersonales</i> 1. Capacidad de trabajo en equipo.	
<i>Sistémicas</i> 1. Aplicar nuevos conocimientos adquiridos. 2. Motivación por la calidad y el logro. 3. Generación de nuevas ideas.	

4. Contenidos

Bloque I: Manipulación de información en bases de datos textuales y la Web.

- Introducción a las aplicaciones inteligentes de la Web
- Crawling y recuperación de información en colecciones de documentos
- Algoritmos para la clasificación de textos
- Algoritmos de agrupación (“clustering”) de documentos

Bloque II: Aplicaciones con informaciones textuales.

- Extracción de información
- Resumen automático
- Minería de opiniones

Bloque III: Tópicos avanzados

- Web Semántica

5. Evaluación del nivel de adquisición de las competencias

Examen teórico al fin del curso (50%), trabajos prácticos (50%). Los trabajos de seminario contribuyen con hasta un 10% de la nota final. Cada parte se aprueba independientemente. Las prácticas se pueden realizar en grupos de hasta 3 personas y se evalúan a partir de una memoria y prueba del software entregado. La teoría se evalúa con un examen escrito al final del trimestre.

6. Bibliografía y recursos didácticos

Bibliografía básica:

Haralambos Marmanis and Dmitry Babenko. Algorithms of the intelligent web. Manning Publications, 2009.

Segaran, T. Programming collective intelligence : building smart web 2.0 applications. O'Reilly, 2007.

Bibliografía complementaria

Baeza-Yates, R. thier Ribeiro-Neto. Modern information retrieval. Reading, Mass. : Addison-Wesley Longman, 1999

Bramer, M. Principles of Data Mining. Springer. 2007.

GATE user guide. Disponible en <http://gate.ac.uk/sale/tao/split.html>

Grigoris Antoniou and Frank van Harmelen. A Semantic Web primer. The MIT Press, 2008.

Jurafsky, Dan & Martin, James. Speech and language processing : an introduction to natural language processing, computational linguistics, and speech recognition, Upper Saddle River, N.J. : Pearson Prentice Hall, 2009

Mani, Inderjeet. Automatic Summarization. John Benjamins Publishing Company. 2001.

Manning, C.D. Prabhakar Raghavan, Hinrich Schütze. Introduction to information retrieval. New York : Cambridge University Press, 2008

Manning, Christopher D. & Schutze, H. Foundations of statistical natural language processing Cambridge, Mass. : MIT Press, 1999

Mitchell, Tom M. Machine learning. New York : McGraw-Hill, 1997.

Pang, B. and Lillian Lee. Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval 2(1-2), 2008.

Pazienza. M.T.. Information extraction: a multidisciplinary approach to an emerging information technology. Springer, 1997.

Ian H. Witten, Eibe Frank. Data mining : practical machine learning tools and techniques. Morgan Kaufman, 2005.

Otros recursos

Artículos científicos serán puestos a disposición de los alumnos a través del aula global.

7. Metodología

La asignatura tiene clases teóricas, seminarios y laboratorios. Los profesores presentarán las bases teóricas exponiendo los contenidos preparados a partir de la bibliografía sugerida y de artículos científicos pertinentes. Los alumnos deberán complementar los conceptos impartidos con lecturas sugeridas por los profesores. En los seminarios los estudiantes presentarán en grupo un tema escogido sobre aplicaciones inteligentes para la Web a partir de las sugerencias de los profesores. En los laboratorios se presentarán enunciados de proyectos a desarrollar y se darán las pautas necesarias para su implementación. Los proyectos tendrán una fecha apropiada para la entrega de los mismos. Para las prácticas se utilizarán herramientas (software) y datos libres.

Bloques de contenido	Horas en el aula			Horas fuera del aula	
	Grupo grande	Grupo mediano	Grupo pequeño		
Bloque I	8	4	0	30	
Bloque II	6	4	2	40	
Bloque III	2	2	2	20	
Examen					
Total	16	10	4		# horas total (ECTS*25)