

PLAN DOCENTE DE LA ASIGNATURA. SEMIPRESENCIAL (Curso de Adaptación).

Curso académico: 2020/2021

Identificación y características de la asignatura			
Código	502286	Créditos ECTS	6
Denominación (español)	Procesamiento Avanzado de la Información		
Denominación (inglés)	Advanced Processing of Information		
Titulaciones	Grado en Información y Documentación		
Centro	Facultad de Ciencias de la Documentación y la Comunicación		
Semestre	2º	Carácter	Obligatorio
Módulo	Tecnologías y Aplicación Práctica de la Información y Documentación		
Materia	Representación y Recuperación de la Información		
Profesor/es			
Nombre	Despacho	Correo-e	Página web
Vicente P. Guerrero Bote	D.03, D.55	guerrero@unex.es	
Área de conocimiento	Biblioteconomía y Documentación		
Departamento	Información y Comunicación		
Profesor coordinador (si hay más de uno)	Vicente P. Guerrero Bote		
Competencias *			
<i>Competencias Básicas</i>			
<ol style="list-style-type: none"> 1. (CB1) Que los estudiantes hayan demostrado poseer y comprender conocimientos en un área de estudio que parte de la base de la educación secundaria general, y se suele encontrar a un nivel que, si bien se apoya en libros de texto avanzados, incluye también algunos aspectos que implican conocimientos procedentes de la vanguardia de su campo de estudio. 2. (CB2) Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio. 3. (CB3) Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética. 4. (CB4) Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado. 5. (CB5) Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía. 			
<i>Competencias Generales</i>			
<ol style="list-style-type: none"> 6. (CG2) Conocimiento de los principios teóricos y metodológicos para la reunión, selección, organización, representación, preservación, recuperación, acceso, difusión e intercambio de la información. 			

* Los apartados relativos a competencias, breve descripción del contenido, actividades formativas, metodologías docentes, resultados de aprendizaje y sistemas de evaluación deben ajustarse a lo recogido en la memoria verificada del título.

<p>7. (CG3) Conocimiento de las tecnologías de la información que se emplean en las unidades y servicios de información.</p> <p>8. (CG4) Habilidades en el manejo de las tecnologías como medio indispensable en los procesos de tratamiento y transferencia de la información.</p>
<p><i>Competencias Específicas</i></p>
<p>9. (CE5) Comprender y aplicar los principios y las técnicas para la reunión, selección, organización, representación, preservación, recuperación, acceso, difusión e intercambio de la información.</p> <p>10. (CE11) Capacidad para autenticar, usar, diseñar y evaluar las fuentes y recursos de información.</p>
<p><i>Competencias Transversales</i></p>
<p>11. (CT3) Habilidades en el uso de Internet y software genérico (ofimática).</p> <p>12. (CT10) Capacidad para el aprendizaje autónomo.</p> <p>13. (CT11) Capacidad para la adaptación a cambios en el entorno.</p> <p>14. (CT12) Capacidad para emprender mejoras y proponer innovaciones.</p>
<p>Contenidos</p>
<p>Breve descripción del contenido*</p>
<p>Algoritmos de clustering, clasificación, condensación documental. Enriquecimiento con la información de enlaces, pagerank.</p>
<p>Temario de la asignatura</p>
<p>Denominación del tema 1: Recuperación de Información. Contenidos del tema 1:</p> <ol style="list-style-type: none"> 1. Historia 2. Modelo de Recuperación de Información. <ol style="list-style-type: none"> i. Modelo Booleano. ii. Modelos de emparejamiento parcial. <ol style="list-style-type: none"> 1. Modelo del Espacio Vectorial. 2. Modelo Probabilístico. 3. Medidas del éxito. <p>Descripción de las actividades prácticas del tema 1: Resolución de Ejemplos en las aulas de informática con la ayuda de una hoja de cálculo y el Campus Virtual.</p>
<p>Denominación del tema 2: Clustering, Clasificación y Condensación Documental. Contenidos del tema 2:</p> <ol style="list-style-type: none"> 1. Clustering y Clasificación. 2. Hipótesis del Clustering. 3. Algoritmos de Clustering. <ol style="list-style-type: none"> i. Distancia/Similitud. ii. Tipos de clustering. <ol style="list-style-type: none"> 1. Algoritmos de clustering jerárquicos. 2. Algoritmos particional de la k-medias. 3. Algoritmos de clustering espectral. 4. Mapas Auto-Organizativos de Kohonen. 4. Mecanismos de condensación documental automatizadas. <ol style="list-style-type: none"> i. Técnicas de Etiquetado. ii. Técnicas de Resumen. <p>Descripción de las actividades prácticas del tema 2: Resolución de Ejemplos en las aulas de informática con la ayuda de una hoja de cálculo y el Campus Virtual.</p>
<p>Denominación del tema 3: Recuperación de Información Web. Contenidos del tema 3:</p> <ol style="list-style-type: none"> 1. Marco de la Recuperación de Información Web. <ol style="list-style-type: none"> i. La Web como Colección Documental <ol style="list-style-type: none"> 1. Tamaño. 2. Dinamismo. 3. Auto-organización.

4. Hiperenlazado.
 - ii. Eficacia en un buscador web.
 - iii. Elementos de un buscador web.
 - iv. Situación en 1998.
2. Algoritmo de PageRank.
3. Algoritmo de HITS.
 - i. Conjunto de resultados.
 - ii. Algoritmo.
 - iii. Algebra lineal subyacente.
 - iv. Pros y contras del algoritmo HITS.
4. Otros mecanismos de ponderación.
5. Fraudes para mejorar de la posición en los buscadores.
 - i. Anteriores a 1998.
 - ii. Granjas de Enlaces (link farming).
 - iii. Google Bombing.
6. Retos actuales de la Recuperación de la Información Web.

Descripción de las actividades prácticas del tema 3: Resolución de Ejemplos en las aulas de informática con la ayuda de una hoja de cálculo y el Campus Virtual.

Actividades formativas*

Horas de trabajo del alumno por tema		Horas teóricas	Actividades prácticas				Actividad de seguimiento	No presencial
Tema	Total	GG	PCH	LAB	ORD	SEM	TP	EP
1	40	1			1		0	38
2	49	1			2		0	46
3	50	0			2		1,5	46,5
Evaluación **	11	1						10
TOTAL	150	3			5		1,5	140,5

GG: Grupo Grande (100 estudiantes).

PCH: prácticas clínicas hospitalarias (7 estudiantes)

LAB: prácticas laboratorio o campo (15 estudiantes)

ORD: prácticas sala ordenador o laboratorio de idiomas (30 estudiantes)

SEM: clases problemas o seminarios o casos prácticos (40 estudiantes).

TP: Tutorías Programadas (seguimiento docente, tipo tutorías ECTS).

EP: Estudio personal, trabajos individuales o en grupo, y lectura de bibliografía.

Metodologías docentes*

- Utilización de material docente en diferentes tipos y formatos.
- Discusión de los contenidos.
- Análisis y resolución de problemas prácticos propuestos.
- Actividades de seguimiento del aprendizaje.
- Autoevaluaciones.
- Actividades experimentales guiadas.

Resultados de aprendizaje*

- Encuadrar el concepto de recuperación de información en del proceso documental.
- Comprender la importancia de la representación documental (indización) y su relación con la recuperación de información.
- Comprender la importancia de una correcta evaluación de la recuperación de la información, y aplicar las técnicas necesarias para llevarla a cabo.

** Indicar el número total de horas de evaluación de esta asignatura.

- Conocer y emplear los distintos modelos de recuperación de información.
- Conocer distintos métodos de clustering aplicados a la documentación.
- Comprender los fundamentos de los métodos de condensación documental automatizadas.
- Conocer el enriquecimiento de la recuperación de información mediante los enlaces web.
- Manejar el algoritmo de PageRank.

Sistemas de evaluación*

1. Evaluación Continua a través de la red. Consistirá en una serie de actividades con las condiciones y plazos que se establezcan a lo largo del curso.
2. Evaluación Final: la evaluación teórica pesará el 60% mientras que los problemas prácticos el 40%.
3. En las recuperaciones que se hagan, las tareas que se proponen a lo largo del curso no serán recuperables.
4. Para cumplir con la normativa de evaluación de la Uex, por defecto se considerará que el alumno opta por el Sistema de Evaluación Continua, a menos que manifieste lo contrario en las tres primeras semanas del semestre a través del Campus Virtual.
5. En caso de optar por la prueba única final el Examen Final constituirá el 100% de la Nota Final. En caso contrario, el Examen Final constituirá el 40% y la Evaluación Continua el 60%. Aunque en todo caso deberá aprobarse el Examen Final para superar la asignatura.

Bibliografía (básica y complementaria)

- Ellis, D. Progress and problems in information retrieval. The Library Association, London, 1996.
- Guerrero-Bote, V.P, Moya-Anegón, F. y Herrero-Solana, V. Document organization using Kohonen's algorithm. Information Processing & Management 38, no. 1 (January): 79-89, 2002.
- Langville, A. N., Meyer, C. D. Google's PageRank and Beyond. Princeton University Press, 2006.
- Moya Anegón, F. Los Sistemas Integrados de Gestión Bibliotecarias. ANABAD, 1994.
- Nooy, W., Mrvar, A., Batagelj, V. Exploratory Social Network Analysis with Pajek. Cambridge University Press, 2005.
- Page, L. Method for node ranking in a linked database. United States Patent Application 2001/6285999, Kind Code B 1, 2001.
- Page, L., Brin, S., Motwani, R. & Winograd, T. The PageRank citation ranking: Bringing order to the Web. Technical report, Stanford University, Stanford, CA, 1998.
- Salton, G. Automatic text procesing: the transformation, analysis, and retrieval of information by computer. Reading, MA: Addison-Wesley, 1989.
- Salton, G. y McGill, M. J. Introduction to modern information retrieval. McGraw-Hill, New York, 1983.

Otros recursos y materiales docentes complementarios

La asignatura cuenta con un aula en el Campus Virtual de la Universidad de Extremadura en la que se encuentran incluidos los principales recursos digitales (temas, presentaciones, cuestionarios, casos prácticos, etc.) para el correcto seguimiento de la misma y realización de la evaluación continua como se indica en el documento de *orientaciones para el estudio* colgado en el correspondiente aula del Campus Virtual.