UNIVERSIDADE DA CORUÑA

Guia docente					
Datos Identificativos					2023/24
Asignatura (*)	Minería de Textos			Código	614544011
Titulación	Máster Universitario en Intelixencia Artificial				
Descriptores					
Ciclo	Periodo	Curso		Tipo	Créditos
Máster Oficial	1º cuatrimestre	Segundo		Optativa	3
Idioma	Inglés				
Modalidad docente	Presencial				
Prerrequisitos					
Departamento					
Coordinador/a	Cabalar Fernandez, Jose Pedro		Correo electrónico	pedro.cabalar@udc.es	
Profesorado	Cabalar Fernandez, Jose Pedro		Correo electrónico	pedro.cabalar@udc.es	
Web					
Descripción general	La asignatura introduce al alumno en la derivación de información y conocimiento a partir del análisis de una colección de documentos en lengua natural, lo que refiere a la práctica totalidad de datos generados y almacenados. Se le capacitará en el análisis del contenido sobre modelos de representación documental enriquecidos, con el fin de abordar aplicaciones concretas sobre distintos dominios. Merecerán especial atención la extracción de información relevante, la determinación de la polaridad contextual (sentimiento) deducible a partir de un contenido y la respuesta automática a preguntas planteadas directamente en lengua natural. Se trata en definitiva de dar respuesta a cuestiones fundamentales en el desarrollo de interfaces, entornos de ayuda a la decisión y acceso a nuevo conocimiento.				

Competencias / Resultados del título	
Código	Competencias / Resultados del título
A2	CE01 - Comprensión y dominio de técnicas para el procesamiento léxico, sintáctico y semántico de textos en lenguaje natural
A3	CE02 - Comprensión y dominio de los fundamentos y técnicas de procesamiento de documentos enlazados, estructurados y no estructurados, y de la representación de su contenido
A4	CE03 - 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
B1	CG01 - 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	CG03 - Buscar y seleccionar la información útil necesaria para resolver problemas complejos, manejando con soltura las fuentes bibliográficas del campo
B4	CG04 - 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
B6	CB01 - 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
B7	CB02 - 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
B10	CB05 - 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
C2	CT02 - Dominar la expresión y la comprensión de forma oral y escrita de un idioma extranjero
C3	CT03 - 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
C7	CT07 - Desarrollar la capacidad de trabajar en equipos interdisciplinares o transdisciplinares, para ofrecer propuestas que contribuyan a un desarrollo sostenible ambiental, económico, político y social
C8	CT08 - Valorar la importancia que tiene la investigación, la innovación y el desarrollo tecnológico en el avance socioeconómico y cultural de la sociedad

Resultados de aprendizaje			
Resultados de aprendizaje	Competencias / Resultados del título		
Conocer, comprender y analizar las técnicas actuales de búsqueda y minería en la web	AM1 AM2 AM3	BM1 BM3 BM4 BM6 BM7 BM10	CM2 CM3 CM7 CM8
Diseñar, implementar y saber usar algoritmos y estructuras de datos para sistemas de recomendación	AM1 AM2	BM1 BM3 BM4 BM6 BM7 BM10	CM2 CM3 CM7 CM8
Saber aplicar diferentes modelos de recuperación y extracción de información, análisis del sentimiento y otras posibles aplicaciones de la minería de textos	AM1 AM2 AM3	BM1 BM3 BM4 BM6 BM7 BM10	CM2 CM3 CM7 CM8

Contenidos	
Tema	Subtema
Análisis de documentos.	Conceptos y definiciones. Estructura argumental, coherencia y co-referencias.
Recuperación y extracción de información	Conceptos y definiciones. Técnicas y herramientas de RI (Recuperación de Informaticón) Técnicas y herramientas de EI (Extracción de Información)
Análisis del sentimiento	Conceptos y definiciones. Técnicas y herramientas. Tendencias actuales.
Búsqueda de respuestas.	Conceptos y definiciones. Arquitecturas típicas, técnicas y herramientas. Tendencias actuales.
Otras aplicaciones de la minería de textos.	Problemas emergentes. Text mining en dominios específicos.

Planificación				
Metodologías / pruebas	Competencias / Resultados	Horas lectivas (presenciales y virtuales)	Horas trabajo autónomo	Horas totales
Prácticas de laboratorio	A2 A3 A4 B1 B3 B4 B6 B7 B10 C2 C3 C7 C8	5	15	20
Trabajos tutelados	A2 A3 A4 B1 B3 B4 B6 B7 B10 C2 C3 C7 C8	5	29	34

Prueba objetiva	A2 A3 A4 B1 B3 B4 B6 B7 B10 C2 C3 C7 C8	1	0	1
Sesión magistral	A2 A3 A4 B1 B3 B4 B6 B7 B10 C2 C3 C7 C8	10	10	20
Atención personalizada		0		0
(*)Los datos que aparecen en la tabla de planificación són de carácter orientativo, considerando la heterogeneidad de los alumnos				

Metodologías	
Metodologías	**Descripción**
Prácticas de laboratorio	Realización de problemas de carácter práctico que incluyen el empleo de herramientas específicas y la programación de software relacionado con los contenidos de la materia. EVALUACION CONTINUA Caracter: Obligatorio Asistencia: No obligatoria EVALUACION GLOBAL Caracter: Obligatorio
Trabajos tutelados	Conjunto de uno o más trabajos teórico-prácticos individuales, entregables y evaluables, sobre los aspectos teóricos presentados en la materia y trabajados en las actividades prácticas desarrolladas por los alumnos. Se trata de una tarea autónoma que contará con la tutorización puntual del profesorado. El resultado se plasmará en una o más memorias con la estructura que se determine. EVALUACION CONTINUA Caracter: Obligatorio Asistencia: No obligatoria EVALUACION GLOBAL Caracter: Obligatorio
Prueba objetiva	Examen escrito
Sesión magistral	Exposición de los contenidos teóricos de la materia. Con el fin de facilitar la comprensión de la misma y aumentar el interés del alumno, se incluirán diversos ejemplos y ejercicios en los que se puede requerir la participación activa del alumno. Se promoverá una actitud activa, fomentando la realización de preguntas y proponiendo cuestione abiertas para la reflexión del alumno,

Atención personalizada	
Metodologías	**Descripción**
Prácticas de laboratorio Trabajos tutelados	El profesor tutorizará al alumno en el laboratorio para la realización de los proyectos que se evaluarán al final de la materia, respondiendo dudas individualmente Seguimiento del trabajo de los alumnos, resolución de dudas generales y puesta en común de problemas específicos de carácter teórico/práctico relacionados con la materia

Evaluación			
Metodologías	**Competencias / Resultados**	**Descripción**	**Calificación**

Prácticas de laboratorio	A2 A3 A4 B1 B3 B4 B6 B7 B10 C2 C3 C7 C8	Evaluación de las prácticas de laboratorio propuestas mediante la entrega de memoria y/o del código desarrollado. La entrega de estas prácticas es obligatoria. Tendrán una fecha de entrega y, opcionalmente, de defensa. - PUNTUACION MINIMA: 4 puntos sobre 10 - RESULTADOS APRENDIZAJE EVALUADOS: RA1	40
Trabajos tutelados	A2 A3 A4 B1 B3 B4 B6 B7 B10 C2 C3 C7 C8	Evaluación de la memoria del trabajo (o trabajos) de investigación tutelado, de caractes teórico-práctico, asignado a cada alumno. Se evaluará la capacidad de síntesis y la completitud y adecuada presentación de las ideas y conceptos relativos al tema escogido. La entrega de estos trabajos es obligatoria. Tendrán una fecha de entrega y, opcionalmente, de defensa. - PUNTUACION MINIMA: 4 puntos sobre 10 - RESULTADOS APRENDIZAJE EVALUADOS: RA1	40
Prueba objetiva	A2 A3 A4 B1 B3 B4 B6 B7 B10 C2 C3 C7 C8	Prueba escrita donde se evaluarán los contenidos y competencias revisados en las sesiones magistrales y los aspectos teóricos de su puesta en práctica llevada a cabo en las sesión prácticas. El tipo de prueba consistirá en un conjunto de preguntas tipo test o cuestiones de respuesta corta sobre conceptos concretos. Se realizará en la fecha oficial prevista en el calendario de la titulación. - PUNTUACION MINIMA: no hay mínimo exigido - RESULTADOS APRENDIZAJE EVALUADOS: RA1	20

Observaciones evaluación

(1) SISTEMA DE EVALUACIÓN CONTÍNUAPRUEBA 1:Prácticas de laboratorioDescripción: Evaluación de la memoria de las prácticas de laboratorio entregadas en las fechas estipuladas.Metodología(s): Prácticas de laboratorio% Calificación: 40%% Mínimo: 4 puntos sobre 10Páxina 4 de 6Competencias evaluadas: A2, B1, C1, C2, C3, D3, D8Resultados aprendizaje evaluados: RA1PRUEBA 2: Trabajo tuteladoDescripción: Evaluación de la memoria del trabajo(s) entregados en las fechas estipuladas y de supresentación en clase.Metodología(s): Trabajo tutelado% Calificación: 40%% Mínimo: 4 puntos sobre 10Competencias evaluadas: A2, A5, B3, B4, C1, C2, C3, D2, D8Resultados aprendizaje evaluados: RA1PRUEBA 3: Examen finalDescripción: Examen tipo test o de conceptos sobre los contenidos teóricos de la materiaMetodología(s): Examen de preguntas objetivas% Calificación: 20%% Mínimo: no hay mínimoCompetencias evaluadas: A1, A2, B1, C1, C2, C3Resultados aprendizaje evaluados: RA1ACLARACIONES ADICIONALESPara superar la materia es preciso alcanzar los mínimos indicados en las pruebas anteriores y sumar en la nota finalponderada un mínimo de 5 puntos sobre 10.En el caso de constatar un comportamiento no ético (copia, plagio) en alguna de las entregas realizadas (total oparcial), se anulará la totalidad de la contribución del correspondiente elemento de evaluación sobre la calificaciónfinal(2) SISTEMA DE EVALUACIÓN GLOBALProcedimiento para la elección de la modalidad de evaluación global:Se asume por defecto la modalidad de evaluación contínua.Los alumnos que opten por la evaluación global deberán comunicarlo via Moovi, mediante los mecanismos que sehabiliten y en el plazo estipulado, una vez superado el plazo de un mes desde el comienzo del cuatrimestrePRUEBA 1: Prácticas de laboratorioDescripción: Evaluación de la memoria de las prácticas de laboratorio entregadas en las fechas estipuladas.Metodología(s): Prácticas de laboratorio% Calificación: 40%% Mínimo: 5 puntos sobre 10Páxina 5 de 6Competencias evaluadas: A2, B1, C1, C2, C3, D3, D8Resultados aprendizaje evaluados: RA1PRUEBA 2: Trabajo tuteladoDescripción: Evaluación de la memoria del trabajo(s) entregados en las fechas estipuladas y de supresentación en clase.Metodología(s): Trabajo tutelado% Calificación: 40%% Mínimo: 5 puntos sobre 10Competencias evaluadas: A2, A5, B3, B4, C1, C2, C3, D2, D8Resultados aprendizaje evaluados: RA1PRUEBA 3: Examen finalDescripción: Examen tipo test o de conceptos sobre los contenidos teóricos de la materiaMetodología(s): Examen de preguntas objetivas% Calificación: 20%% Mínimo: 5 puntos sobre 10Competencias evaluadas: A1, A2, B1, C1, C2, C3Resultados aprendizaje evaluados: RA1ACLARACIONES ADICIONALESPara superar la materia es preciso alcanzar los mínimos indicados en las pruebas anteriores y sumar en la nota finalponderada un mínimo de 5 puntos sobre 10.En el caso de constatar un comportamiento no ético (copia, plagio) en alguna de las entregas realizadas (total oparcial), se anulará la totalidad de la contribución del correspondiente elemento de evaluación sobre la calificaciónfinal(3) CRITERIOS DE EVALUACIÓN PARA CONVOCATORIA EXTRAORDINARIA Y FIN DE CARRERASe emplearán los sistemas de evaluación continua y global expuestos anteriormente.En estas convocatorias, los alumnos sólo deberán realizar las pruebas en las cuales no hubieran obtenido la calificaciónmínima indicada.(4) PROCESO DE CALIFICACIÓN DE ACTASEn el caso de los alumnos que superen parte de los elementos evaluados, pero no alcancen el mínimo preciso para aprobarla materia completa, la calificación a incluir en las respectivas actas se calculará como el mínimo entre el promedioponderado de las partes superadas y 4,9.(5) FECHAS DE EVALUACIÓNEl calendario de pruebas de evaluación aprobado oficialmente por la Junta de Centro de la ESEI se encuentra publicado en lapágina web http://www.esei.uvigo.es(6) EMPLEO DE DISPOSITIVOS MOVILESSe recuerda a todo el alumnado la prohibición del uso de dispositivos móviles en ejercicios y prácticas, en cumplimiento delPáxina 6 de 6artículo 13.2.d) del Estatuto del Estudiante Universitario, relativo a los deberes del estudiantado universitario, que estableceel deber de "Abstenerse de la utilización o cooperación en procedimientos fraudulentos en las pruebas de evaluación, en lostrabajos que se realicen o en documentos oficiales de la universidad."(7) CONSULTA/SOLICITUD DE TUTORÍASLas tutorías pueden consultarse a través de la página personal del profesorado, accesible a través dehttps://esei.uvigo.es/docencia/profesorado/

Fuentes de información	
Básica	- Berry, M. W., & Kogan, J. (Eds.)., (2010). Text mining: applications and theory. John Wiley & Sons - Taeho Jo (2019). Text Mining: Concepts, Implementation, and Big Data Challenge (Studies in Big Data Book 45). Springer
Complementária	

Recomendaciones
Asignaturas que se recomienda haber cursado previamente
Comprensión del Lenguaje Natural/614544008 Modelado del Lenguaje/614544009 Inteligencia Web y Tecnologías Semánticas/614544010
Asignaturas que se recomienda cursar simultáneamente

Asignaturas que continúan el temario
Otros comentarios

(*) La Guía Docente es el documento donde se visualiza la propuesta académica de la UDC. Este documento es público y no se puede modificar, salvo cosas excepcionales bajo la revisión del órgano competente de acuerdo a la normativa vigente que establece el proceso de elaboración de guías