



Guía Docente

Guía Docente				
Datos Identificativos				2023/24
Asignatura (*)	Minería de Textos		Código	614544011
Titulación	Máster Universitario en Intelixencia Artificial			
Descriptor				
Ciclo	Período	Curso	Tipo	Créditos
Mestrado Oficial	1º cuatrimestre	Segundo	Optativa	3
Idioma	Inglés			
Modalidade docente	Presencial			
Prerrequisitos				
Departamento				
Coordinación	Cabalar Fernandez, Jose Pedro	Correo electrónico	pedro.cabalar@udc.es	
Profesorado	Cabalar Fernandez, Jose Pedro	Correo electrónico	pedro.cabalar@udc.es	
Web				
Descrición xeral	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 do título

Código	Competencias do título
A2	CE01 - Comprensión e dominio de técnicas para o procesamento léxico, sintáctico e semántico de textos en linguaxe natural
A3	CE02 - Comprensión e dominio dos fundamentos e técnicas de procesamento de documentos enlazados, estruturados e non estruturados, e da representación do seu contido
A4	CE03 - Comprensión e coñecemento das técnicas de representación e procesado de coñecemento mediante ontoloxías, grafos e RDF, así como das ferramentas asociadas ás mesmas
B1	CG01 - Manter e extender os plantexamentos teóricos fundados para permitir a introducción e explotación de tecnoloxías novas e avanzadas no eido da Intelixencia Artificial
B3	CG03 - Buscar e seleccionar a información útil necesaria para resolver problemas complexos, manexando con soltura as fontes bibliográficas do campo
B4	CG04 - Elaborar axeitadamente e con certa orixinalidade composicións escritas ou argumentos motivados, redactar plans, proxectos de traballo, artigos científicos e formular hipóteses razoables no campo
B6	CB01 - Poseer e comprender coñecementos que aporten unha base ou oportunidade de ser orixináis no desenvolvemento e/ou aplicación de ideas, a miúdo nun contexto de investigación
B7	CB02 - Que os estudantes saiban aplicar os coñecementos adquiridos e posúan capacidade de resolución de problemas en entornos novos ou pouco coñecidos dentro de contextos máis amplos (ou multidisciplinares) relacioados coa su área de estudo
B10	CB05 - Que os estudantes posúan as habilidades de aprendizaxe que lles permitan continuar estudando dun xeito que terá que ser en gran medida autodirixido ou autónomo
C2	CT02 - Dominar a expresión e comprensión, de xeito oral e escrito, dun idioma estranxeiro
C3	CT03 - 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
C7	CT07 - Desenvolver a capacidade de traballar en equipos interdisciplinares ou transdisciplinares, para ofrecer propostas que contribúan a un desenrolo sostible ambiental, económico, político e social
C8	CT08 - Valorar a importancia que ten a investigación, a innovación e o desenrolo tecnolóxico no avance socioeconómico e cultural da sociedade



Resultados da aprendizaxe

Resultados de aprendizaxe	Competencias do título		
Coñecer, comprender e analizar as técnicas actuais de búsqueda e minería na web	AM1 AM2 AM3	BM1 BM3 BM4 BM6 BM7 BM10	CM2 CM3 CM7 CM8
Deseñar, implementar e saber usar algoritmos e estruturas de datos para sistemas de recomendación.	AM1 AM2	BM1 BM3 BM4 BM6 BM7 BM10	CM2 CM3 CM7 CM8
Saber aplicar diferentes modelos de recuperación e extracción de información, análise do sentimento e outras posibles aplicacións da minería de textos	AM1 AM2 AM3	BM1 BM3 BM4 BM6 BM7 BM10	CM2 CM3 CM7 CM8

Contidos

Temas	Subtemas
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 Información) 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

Metodoloxías / probas	Competencias	Horas presenciais	Horas non presenciais / traballo autónomo	Horas totais
Prácticas de laboratorio	A2 A3 A4 B1 B3 B4 B6 B7 B10 C2 C3 C7 C8	5	15	20
Traballos tutelados	A2 A3 A4 B1 B3 B4 B6 B7 B10 C2 C3 C7 C8	5	29	34



Proba obxectiva	A2 A3 A4 B1 B3 B4 B6 B7 B10 C2 C3 C7 C8	1	0	1
Sesión maxistral	A2 A3 A4 B1 B3 B4 B6 B7 B10 C2 C3 C7 C8	10	10	20
Atención personalizada		0		0
*Os datos que aparecen na táboa de planificación son de carácter orientativo, considerando a heteroxeneidade do alumnado				

Metodoloxías	
Metodoloxías	Descrición
Prácticas de laboratorio	Realización de problemas de carácter práctico que inclúen 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
Traballos 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
Proba obxectiva	Examen escrito
Sesión maxistral	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 cuestiones abiertas para la reflexión del alumno,

Atención personalizada	
Metodoloxías	Descrición
Prácticas de laboratorio Traballos 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

Avaliación			
Metodoloxías	Competencias	Descrición	Cualificación



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
Traballos 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 caracters 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
Proba obxectiva	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

Observación avaliación



(1) SISTEMA DE EVALUACIÓN CONTÍNUA PRUEBA 1: Prácticas de laboratorio Descripción: 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 10 Páxina 4 de 6 Competencias evaluadas: A2, B1, C1, C2, C3, D3, D8 Resultados aprendizaje evaluados: RA1 PRUEBA 2: Trabajo tutelado Descripción: 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 10 Competencias evaluadas: A2, A5, B3, B4, C1, C2, C3, D2, D8 Resultados aprendizaje evaluados: RA1 PRUEBA 3: Examen final Descripción: Examen tipo test o de conceptos sobre los contenidos teóricos de la materia Metodología(s): Examen de preguntas objetivas % Calificación: 20% % Mínimo: no hay mínimo Competencias evaluadas: A1, A2, B1, C1, C2, C3 Resultados aprendizaje evaluados: RA1 ACLARACIONES ADICIONALES Para superar la materia es preciso alcanzar los mínimos indicados en las pruebas anteriores y sumar en la nota final ponderada 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 o parcial), se anulará la totalidad de la contribución del correspondiente elemento de evaluación sobre la calificación final.

(2) SISTEMA DE EVALUACIÓN GLOBAL Procedimiento para la elección de la modalidad de evaluación global: Se asume por defecto la modalidad de evaluación continua. Los alumnos que opten por la evaluación global deberán comunicarlo via Moovi, mediante los mecanismos que se habiliten y en el plazo estipulado, una vez superado el plazo de un mes desde el comienzo del cuatrimestre PRUEBA 1: Prácticas de laboratorio Descripción: 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 10 Páxina 5 de 6 Competencias evaluadas: A2, B1, C1, C2, C3, D3, D8 Resultados aprendizaje evaluados: RA1 PRUEBA 2: Trabajo tutelado Descripción: 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 10 Competencias evaluadas: A2, A5, B3, B4, C1, C2, C3, D2, D8 Resultados aprendizaje evaluados: RA1 PRUEBA 3: Examen final Descripción: Examen tipo test o de conceptos sobre los contenidos teóricos de la materia Metodología(s): Examen de preguntas objetivas % Calificación: 20% % Mínimo: 5 puntos sobre 10 Competencias evaluadas: A1, A2, B1, C1, C2, C3 Resultados aprendizaje evaluados: RA1 ACLARACIONES ADICIONALES Para superar la materia es preciso alcanzar los mínimos indicados en las pruebas anteriores y sumar en la nota final ponderada 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 o parcial), se anulará la totalidad de la contribución del correspondiente elemento de evaluación sobre la calificación final.

(3) CRITERIOS DE EVALUACIÓN PARA CONVOCATORIA EXTRAORDINARIA Y FIN DE CARRERA Se 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 calificación mínima indicada.

(4) PROCESO DE CALIFICACIÓN DE ACTAS En 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 promedio ponderado de las partes superadas y 4,9.

(5) FECHAS DE EVALUACIÓN El calendario de pruebas de evaluación aprobado oficialmente por la Junta de Centro de la ESEI se encuentra publicado en la página web <http://www.esei.uvigo.es> (6) EMPLEO DE DISPOSITIVOS MÓVILES Se recuerda a todo el alumnado la prohibición del uso de dispositivos móviles en ejercicios y prácticas, en cumplimiento del artículo 13.2.d) del Estatuto del Estudiante Universitario, relativo a los deberes del estudiantado universitario, que establece el deber de "Abstenerse de la utilización o cooperación en procedimientos fraudulentos en las pruebas de evaluación, en los trabajos que se realicen o en documentos oficiales de la universidad."

(7) CONSULTA/SOLICITUD DE TUTORÍAS Las tutorías pueden consultarse a través de la página personal del profesorado, accesible a través de <https://esei.uvigo.es/docencia/profesorado/>

Fontes de información

Bibliografía básica	- Berry, M. W., & Kogan, J. (Eds.), (2010). Text mining: applications and theory. John Wiley & Sons - Taehoon Jo (2019). Text Mining: Concepts, Implementation, and Big Data Challenge (Studies in Big Data Book 45). Springer
Bibliografía complementaria	

Recomendacións

Materias que se recomienda ter cursado previamente

Comprensión da Linguaxe Natural/614544008

Modelado da Linguaxe/614544009

Intelixencia Web e Tecnoloxías Semánticas/614544010

Materias que se recomienda cursar simultaneamente



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