

Facultad de Ciencias

GUÍA DOCENTE DE LA ASIGNATURA

M1971 - Semántica, Datos Conectados y Minería de Datos Textual

Máster Universitario en Ciencia de Datos / Master in Data Science
Optativa. Curso 1

Curso Académico 2019-2020

1. DATOS IDENTIFICATIVOS

Título/s	Máster Universitario en Ciencia de Datos / Master in Data Science	Tipología y Curso	Optativa. Curso 1
Centro	Facultad de Ciencias		
Módulo / materia	ESPECIALIZACION INTELIGENCIA EN CIENCIA DE DATOS		
Código y denominación	M1971 - Semántica, Datos Conectados y Minería de Datos Textual		
Créditos ECTS	4	Cuatrimestre	Cuatrimestral (2)
Web			
Idioma de impartición	Español	English friendly	Sí
		Forma de impartición	Presencial

Departamento	DPTO. MATEMATICAS, ESTADISTICA Y COMPUTACION
Profesor responsable	CRISTINA TIRNAUCA
E-mail	cristina.tirnauca@unican.es
Número despacho	Facultad de Ciencias. Planta: + 1. DESPACHO PROFESORES (1046)
Otros profesores	JOSE JAVIER RAMASCO SUQUIA MIGUEL EXPOSITO MARTIN FERNANDO AGUILAR GOMEZ

2. CONOCIMIENTOS PREVIOS

Los conocimientos adquiridos en las otras asignaturas de la misma materia.

3. COMPETENCIAS GENÉRICAS Y ESPECÍFICAS DEL PLAN DE ESTUDIOS TRABAJADAS

Competencias Genéricas
Integrarse eficazmente en un grupo de trabajo y trabajar en equipo, compartir la información disponible e integrar su actividad en la actividad del grupo colaborando de forma activa en la consecución de objetivos comunes.
Capacidad de estudio, síntesis y autonomía suficientes para desarrollar de forma autónoma proyectos básicos de investigación.
Saber preparar y conducir presentaciones, ante públicos especializado, sobre una investigación o proyecto científico.
Conocer las herramientas metodológicas necesarias para desarrollar proyectos avanzados.
Capacidad de actualización de los conocimientos expuestos en el ámbito de la comunidad científica.
Competencias Específicas
Utilizar el análisis predictivo para analizar grandes volúmenes de datos y descubrir nuevas relaciones.
Utilizar técnicas estadísticas apropiadas sobre los datos disponibles para lograr una visión adecuada de los mismos.
Desarrollar y aplicar soluciones computacionales para problemas en un cierto dominio de aplicación, usando una amplia gama de plataformas de análisis de datos.
Aplicar el ingenio propio para resolver problemas complejos y desarrollar ideas innovadoras.
Competencias Básicas
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.
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.
Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.
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.
Competencias Transversales
Analizar y combinar información utilizando diferentes fuentes.
Conocer la problemática ética y legal relacionada con el análisis de datos y entender su importancia para una sociedad basada en los valores de la libertad, la justicia, la igualdad y el pluralismo.
Dominio de la gestión del tiempo.
Afrontar tareas y situaciones críticas.
Capacidad de trabajo autónomo y toma de decisiones.
Capacidades asociadas al trabajo en equipo: cooperación, liderazgo, saber escuchar.

3.1 RESULTADOS DE APRENDIZAJE

- Entender métodos complementarios para el análisis de grandes masas de datos no estructurados, introduciéndose en el campo de la minería de textos (y minería Web)
- Entender los fundamentos de la representación y análisis de datos con redes complejas

4. OBJETIVOS

El principal objetivo de esta asignatura es que el estudiante sepa aplicar los conocimientos básicos sobre las distintas metodologías y técnicas de aprendizaje automático de forma crítica en problemas reales, incluyendo la minería de textos y la minería Web. Un segundo objetivo, de tipo práctico, es proporcionar al alumno las capacidades y herramientas estándar necesarias para poder llevar a cabo de manera autónoma proyectos de analítica de datos.

5. MODALIDADES ORGANIZATIVAS Y MÉTODOS DOCENTES

ACTIVIDADES	HORAS DE LA ASIGNATURA
ACTIVIDADES PRESENCIALES	
HORAS DE CLASE (A)	
- Teoría (TE)	10
- Prácticas en Aula (PA)	10
- Prácticas de Laboratorio (PL)	10
- Horas Clínicas (CL)	
Subtotal horas de clase	30
ACTIVIDADES DE SEGUIMIENTO (B)	
- Tutorías (TU)	7
- Evaluación (EV)	3
Subtotal actividades de seguimiento	10
Total actividades presenciales (A+B)	40
ACTIVIDADES NO PRESENCIALES	
Trabajo en grupo (TG)	30
Trabajo autónomo (TA)	30
Tutorías No Presenciales (TU-NP)	
Evaluación No Presencial (EV-NP)	
Total actividades no presenciales	60
HORAS TOTALES	100

6. ORGANIZACIÓN DOCENTE

CONTENIDOS		TE	PA	PL	CL	TU	EV	TG	TA	TU-NP	EV-NP	Semana
1	Redes semánticas	2,00	2,00	2,00	0,00	1,00	0,00	5,00	5,00	0,00	0,00	1
2	Ontologías y aprendizaje de ontologías	2,00	4,00	4,00	0,00	2,00	0,00	10,00	10,00	0,00	0,00	2
3	Linked data	2,00	2,00	2,00	0,00	1,00	0,00	5,00	5,00	0,00	0,00	3
4	Análisis de redes complejas	2,00	2,00	0,00	0,00	1,00	0,00	5,00	5,00	0,00	0,00	4
5	Minería de textos y minería Web	2,00	0,00	2,00	0,00	2,00	3,00	5,00	5,00	0,00	0,00	5
TOTAL DE HORAS		10,00	10,00	10,00	0,00	7,00	3,00	30,00	30,00	0,00	0,00	

Esta organización tiene carácter orientativo.

TE	Horas de teoría
PA	Horas de prácticas en aula
PL	Horas de prácticas de laboratorio
CL	Horas Clínicas
TU	Horas de tutoría
EV	Horas de evaluación
TG	Horas de trabajo en grupo
TA	Horas de trabajo autónomo
TU-NP	Tutorías No Presenciales
EV-NP	Evaluación No Presencial

7. MÉTODOS DE LA EVALUACIÓN

Descripción	Tipología	Eval. Final	Recuper.	%
Valoración de informes y trabajos escritos	Trabajo	No	Sí	50,00
Calif. mínima	0,00			
Duración				
Fecha realización	Durante el cuatrimestre			
Condiciones recuperación				
Observaciones	Si el número de participantes lo permite, el trabajo se presentará oralmente una vez finalizadas las clases presenciales de la asignatura. Si el trabajo se presenta fuera del plazo establecido por el profesor, la calificación máxima en este apartado será un 8 sobre 10.			
Examen final	Examen escrito	Sí	No	30,00
Calif. mínima	0,00			
Duración				
Fecha realización	Al final del cuatrimestre			
Condiciones recuperación				
Observaciones				
Seguimiento de actividades presenciales	Actividad de evaluación con soporte virtual	No	No	20,00
Calif. mínima	0,00			
Duración				
Fecha realización	Durante el cuatrimestre			
Condiciones recuperación				
Observaciones				
TOTAL				100,00
Observaciones				
Observaciones para alumnos a tiempo parcial				
Los estudiantes a tiempo parcial tienen dos opciones: seguir la evaluación continua (podrán responder a las preguntas a través de moodle) o realizar un proyecto (igual que sus compañeros) que tendrán que presentar y defender oralmente si el número de participantes lo permite. Si deciden no seguir la evaluación continua, cambiarían los porcentajes: 60% trabajo escrito y 40% examen final.				

8. BIBLIOGRAFÍA Y MATERIALES DIDÁCTICOS

BÁSICA
Sholom M. Weiss, Nitin Indurkha, Tong Zhang, Fred Damerau. Text mining: predictive methods for analyzing unstructured information. Springer Science+Business Media (2005)
Juan Antonio Pastor Sánchez. Tecnologías de la web semántica. UOC (2012)
Toby Segaran, Colin Evans, Jamie Taylor. Programming the semantic web. O'Reilly (2009)
Complementaria
Jiawei Han, Micheline Kamber. Data mining: concepts and techniques. Morgan Kaufmann Publishers (2006).

9. SOFTWARE

PROGRAMA / APLICACIÓN	CENTRO	PLANTA	SALA	HORARIO
R y/o Python				

10. COMPETENCIAS LINGÜÍSTICAS

- Comprensión escrita Comprensión oral
 Expresión escrita Expresión oral
 Asignatura íntegramente desarrollada en inglés

Observaciones