



Guía docente

Datos Identificativos					2017/18
Asignatura (*)	Inteligencia computacional para datos de alta dimensionalidad	Código	614522024		
Titulación	Mestrado Universitario en Bioinformática para Ciencias da Saúde				
Descritores					
Ciclo	Periodo	Curso	Tipo	Créditos	
Máster Oficial	1º cuatrimestre	Segundo	Obligatoria	3	
Idioma					
Modalidad docente	Presencial				
Prerrequisitos					
Departamento	Computación				
Coordinador/a		Correo electrónico			
Profesorado	Alonso Betanzos, Maria Amparo	Correo electrónico	amparo.alonso.betanzos@udc.es		
Web					
Descripción general					

Competencias / Resultados del título

Código	Competencias / Resultados del título

Resultados de aprendizaje

Resultados de aprendizaje	Competencias / Resultados del título		
Conocer y comprender los paradigmas y aspectos más relevantes del tratamiento de bases de datos de alta dimensión	A1	B1	C1
	A1	B1	C1
	A1	B1	C1
	A1		C1
	A1		
Conocer y saber aplicar los principales métodos de minería de datos; conocer las plataformas y los paradigmas principales que se emplean en el campo.	A1	B1	C1
	A1	B1	C1
	A1	B1	C1
	A1		C1
	A1		

Contenidos

Tema	Subtema
Introducción a Big data.	Qué es Big Data Principales características del Big data Principales campos de aplicación
Minería de datos y alta dimensión	Analítica Big data Técnicas de preprocesado MapReduce



Modelos de programación Batch	Hadoop Resilient Distributed datasets Programación batch en Spark
Modelos de programación streaming	Conceptos básicos Kafka, Apache Storm, Spark streaming

Planificación				
Metodologías / pruebas	Competencias / Resultados	Horas lectivas (presenciales y virtuales)	Horas trabajo autónomo	Horas totales
Sesión magistral	A4 A12 A14 A21 A43 A45 C1 C2 C6	7	14	21
Solución de problemas	A25 A33 A41 B1 B6 C3	8	16	24
Trabajos tutelados	A21 B3 B6 C1 C2 C3 C6	4	4	8
Seminario	A21 B1 B3 B6	4	4	8
Prueba mixta	A4 A12 B1 B3 C1 C3 C6	4	10	14
Atención personalizada		0	0	0

(*) Los datos que aparecen en la tabla de planificación són de carácter orientativo, considerando la heterogeneidad de los alumnos

Metodologías	
Metodologías	Descripción
Sesión magistral	Utilizada durante las clases presenciales teóricas para exponer el núcleo básico de conocimientos que luego los alumnos tendrán que saber utilizar y ampliar en las prácticas, seminarios y trabajos del curso
Solución de problemas	Utilización de técnicas de minería de datos en alta dimensión. Uso de paradigmas Big data Realización de una práctica en una plataforma específica de Big data
Trabajos tutelados	Entrega de un breve trabajo que se discutirá en clase sobre algún aspecto concreto de la asignatura.
Seminario	Exposición de un trabajo específico de investigación que involucre tecnologías de alta dimensionalidad
Prueba mixta	Se realizará al final del cuatrimestre sobre los contenidos tratados a lo largo del curso.

Atención personalizada	
Metodologías	Descripción
Seminario	En el esquema de carácter práctico utilizado en esta asignatura, las tutorías resultan un recurso fundamental muy utilizado por los alumnos, sobre todo debido a la complejidad de algunos conceptos de la asignatura, en función de las titulaciones de entrada de los diferentes alumnos. Los alumnos pueden realizar dos tipos de tutorías: virtuales y presenciales. Las primeras pueden utilizarlas para realizar dudas muy concretas de respuesta rápida. Las más comunes se irán depositando en un apartado de ?Preguntas Frecuentes? que deberán consultar antes de enviar una nueva pregunta.
Solución de problemas	
Trabajos tutelados	
Prueba mixta	
Sesión magistral	

Evaluación			
Metodologías	Competencias / Resultados	Descripción	Calificación
Seminario	A21 B1 B3 B6		0



Trabajos tutelados	A21 B3 B6 C1 C2 C3 C6	Nota correspondiente a la parte práctica de la asignatura, que comprende tanto los desarrollos realizados sobre las plataformas, como los trabajos entregados.	50
Prueba mixta	A4 A12 B1 B3 C1 C3 C6	Se realizará una prueba con cuestiones relativas a las partes teóricas de la asignatura	50
Sesión magistral	A4 A12 A14 A21 A43 A45 C1 C2 C6		0

Observaciones evaluación

Fuentes de información

Básica	<ul style="list-style-type: none">- Vladimir Bacovski. (2015). Introduction to Big Data An Overview of Fundamental Big Data Concepts, Tools, Techniques and Practices.. O'Reilly Media- Venkat Ankam (2016.). Big Data Analytics. Packt Publishing- Tom White (2015). Hadoop: The Definitive Guide. O'Reilly Media- Thilina Gunarathne (2015). Hadoop MapReduce v2 Cookbook. Packt Publishing- Holden Karau, Andy Konwinski, Patrick Wendell, Matei Zaharia (2015). Learning Spark. O'Reilly Media- Sean T. Allen, Matthew Jankowski, and Peter Pathirana (2015). Storm Applied. . O'Reilly Media
Complementaria	

Recomendaciones

Asignaturas que se recomienda haber cursado previamente

Inteligencia computacional para bioinformática/614522012

Métodos estadísticos avanzados en bioinformática/614522009

Computación de altas prestaciones en bioinformática/614522011

Introducción a la programación/614522001

Fundamentos de inteligencia artificial/614522003

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