UNIVERSIDADE DA CORUÑA

Guia docente					
Datos Identificativos					2021/22
Asignatura (*)	Inteligencia computacional para datos de alta dimensionalidad		Código		614522024
Titulación	Mestrado Universitario en Bioinformática para Ciencias da Saúde				
Descriptores					
Ciclo	Periodo	Curso	Tipo	Créditos	
Máster Oficial	1º cuatrimestre	Segundo	Obligatoria	3	
Idioma	CastellanoGallegoInglés				
Modalidad docente	Presencial				
Prerrequisitos					
Departamento	Ciencias da Computación e Tecnoloxías da InformaciónComputación				
Coordinador/a	Eiras Franco, Carlos	Correo electrónico	carlos.eiras.franco@udc.es		
Profesorado	Cancela Barizo, Brais	Correo electrónico	brais.cancela@udc.es		
Web	moodle.udc.es				
Descripción general	En esta materia se trabajará en los fundamentos y aplicación práctica de las bases de datos de alta dimensión y en la aplicación de técnicas de minería de datos en el ámbito de la bioinformática				
Plan de contingencia	1. Modificaciones en los contenidos 2. Metodologías *Metodologías docentes que se mantienen *Metodologías docentes que se modifican 3. Mecanismos de atención personalizada al alumnado 4. Modificacines en la evaluación *Observaciones de evaluación: 5. Modificaciones de la bibliografía o webgrafía				

Competencias / Resultados del título	
Código	Competencias / Resultados del título
A2	CE2 - Definir, evaluar y seleccionar la arquitectura y el software más adecuado para la resolución de un problema en el campo de la Bioinformática
A3	CE3 - Analizar, diseñar, desarrollar, implementar, verificar y documentar soluciones software eficientes sobre la base de un conocimiento adecuado de las teorías, modelos y técnicas actuales en el campo de la Bioinformática
A4	CE4 - Capacidad para adquirir, obtener, formalizar y representar el conocimiento humano en una forma computable para la resolución de problemas mediante un sistema informático en cualquier ámbito de aplicación, particularmente los relacionados con aspectos de computación, percepción y actuación en aplicaciones Bioinformáticas
A6	CE6 - Capacidad para identificar las herramientas software y fuentes de datos de bioinformática más relevantes, y adquirir destreza en su uso
B1	CB6 - 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
B2	CB7 - 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

B3	CB8 - Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios
B6	CG1 - Buscar y seleccionar la información útil necesaria para resolver problemas complejos, manejando con soltura las fuentes bibliográficas del campo
B7	CG2 - Mantener y extender planteamientos teóricos fundados para permitir la introducción y explotación de tecnologías nuevas y avanzadas
C1	CT1 - Expresarse correctamente, tanto de forma oral como escrita, en las lenguas oficiales de la comunidad autónoma
C3	CT3 - 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
C6	CT6 - Valorar críticamente el conocimiento, la tecnología y la información disponible para resolver los problemas con los que deben enfrentarse

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	AP2 AP3 AP4 AP6	BP1 BP2 BP3 BP6 BP7	CP1 CP3 CP6
Conocer y saber aplicar los principales métodos de minería de datos; conocer las plataformas y los paradigmas principales que se emplean en el campo.	AP2 AP3 AP4 AP6	BP1 BP2 BP3 BP6 BP7	CP1 CP3 CP6

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 C1 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	A2 A3 A4 A6 B1 B2 B3 B6 B7 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 Solución de problemas Trabajos tutelados Prueba mixta Sesión magistral	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.

Evaluación			
Metodologías	**Competencias / Resultados**	**Descripción**	**Calificación**
Seminario	A21 B1 B3 B6	Seminarios de temas específicos	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	A2 A3 A4 A6 B1 B2 B3 B6 B7 C1 C3 C6	Se realizará una prueba con cuestiones relativas a las partes teóricas de la asignatura	50
Sesión magistral	A4 C1 C6	Clases presenciales	0

Observaciones evaluación

Fuentes de información

Básica	- Vladimir Bacvanski. (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
Complementária	

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