



## Guía docente

Datos Identificativos					2022/23
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	Eiras Franco, Carlos	Correo electrónico	carlos.eiras.franco@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				

## 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	12	24	36
Trabajos tutelados	A2 A3 A4 A6 B3 B6 C1 C3	8	24	32
Prueba mixta	A2 A3 A4 A6 B1 B2 B3 B6 B7 C1 C3 C6	2	4	6
Atención personalizada		1	0	1

(\*) 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.
Trabajos tutelados	Elaboración y entrega de trabajos en los que se apliquen las tecnologías y técnicas vistas en la teoría.
Prueba mixta	Se realizará al final del cuatrimestre sobre los contenidos tratados a lo largo del curso.

Atención personalizada	
Metodologías	Descripción



Trabajos tutelados Prueba mixta Sesión magistral	<p>Las tutorías se consideran una parte importante dentro del desarrollo de la asignatura. Están orientadas de tal manera que las/los estudiantes tengan y/o puedan consultar distintas cuestiones como:</p> <ol style="list-style-type: none"> <li>1. Posibilidades de desarrollo profesional</li> <li>2. Problemas en el desarrollo de las prácticas</li> <li>3. Maneras de enfocar/organizar las prácticas</li> <li>4. Resolución de dudas sobre las cuestiones teóricas.</li> </ol> <p>La resolución de dudas y cuestiones se hará en las horas de clase o en las horas establecidas como tutorías de cada profesor.</p>
--	--

Evaluación			
Metodologías	Competencias / Resultados	Descripción	Calificación
Trabajos tutelados	A2 A3 A4 A6 B3 B6 C1 C3	Nota correspondiente a la parte práctica de la asignatura, que comprende los trabajos entregados.	80
Prueba mixta	A2 A3 A4 A6 B1 B2 B3 B6 B7 C1 C3 C6	Se realizará una prueba con cuestiones relativas tanto a las partes teóricas de la asignatura como a los trabajos entregados	20

Observaciones evaluación
<p>La/el estudiante deberá obtener una nota mínima de 4 sobre 10 puntos en la prueba mixta y una nota mínima de 4 sobre 10 puntos en los trabajos tutelados. De no alcanzar esta nota mínima en alguno de dichos apartados, la nota de la materia será la menor de las dos. Cumpliendo este requisito, será necesario alcanzar una nota igual o superior a 5 en el global de la asignatura para superarla. Un/una estudiante se considerará presentado/a en una convocatoria si hace entrega de algún trabajo tutelado o si se presenta a la prueba mixta. El trabajo entregado deberá ser original del/de la estudiante. De acuerdo al artículo 14, apartado 4, de la normativa*, la entrega de trabajos no originales o con partes duplicadas (sea por copias entre compañeros o por obtención de otras fuentes...) llevará una nota global de SUSPENSO en la convocatoria correspondiente, tanto para el/la estudiante que presente material copiado como a quien lo haya facilitado, invalidando cualquier otra calificación obtenida en las actividades evaluables.*</p> <p>Normativa de evaluación, revisión y reclamación de las calificaciones de los estudios de grado y máster universitario, aprobada por el Consejo de Gobierno de la Universidade da Coruña el 19 de diciembre de 2013. Sobre la responsabilidad compartida de los trabajos en grupo. En las actividades que se llevan a cabo en grupos, tales como los trabajos tutelados, todos los miembros del grupo serán responsables solidarios del trabajo realizado y entregado, así como de las consecuencias que se deriven del incumplimiento de las normas de autoría del mismo. Segunda oportunidad y convocatorias posteriores En la segunda oportunidad, se mantiene la nota obtenida en los trabajos tutelados. Los alumnos que se presenten a esta segunda oportunidad deberán realizar la prueba mixta. Opcionalmente, y solo en el caso de no haber presentado los trabajos tutelados en la primera oportunidad o que estas hubiesen obtenido la calificación de SUSPENSO (calificación menor que 5), se podrán presentar nuevamente los trabajos tutelados en la segunda oportunidad. En caso de suspender la asignatura, los trabajos tutelados con nota igual o superior a 5 se podrán guardar para cursos posteriores con calificación de aprobado (5) si el alumno/a así lo desea. Para cada trabajo, el/la estudiante tendrá la opción de entregar un nuevo trabajo tutelado que sustituirá la nota de la anterior. Los trabajos no se guardarán más de un curso. Matrícula con dispensa académica Los/las estudiantes con matrícula con dispensa académica deberán entregar los trabajos tutelados en las fechas establecidas. Es responsabilidad de dichos/as estudiantes poner en conocimiento del profesor su circunstancia.</p>

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



## 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