



Guía docente

| Datos Identificativos | | | | | 2024/25 |
|-----------------------|--|--------------------|-----------------------------|----------|---------|
| Asignatura (*) | Ingeniería de Datos | Código | 614544002 | | |
| Titulación | Máster Universitario en Intelixencia Artificial | | | | |
| Descriptorios | | | | | |
| Ciclo | Periodo | Curso | Tipo | Créditos | |
| Máster Oficial | 1º cuatrimestre | Primero | Obligatoria | 3 | |
| Idioma | Inglés | | | | |
| Modalidad docente | Presencial | | | | |
| Prerrequisitos | | | | | |
| Departamento | Ciencias da Computación e Tecnoloxías da Información | | | | |
| Coordinador/a | Bernardo Roca, Guillermo de | Correo electrónico | guillermo.debernardo@udc.es | | |
| Profesorado | Bernardo Roca, Guillermo de | Correo electrónico | guillermo.debernardo@udc.es | | |
| Web | | | | | |
| Descripción general | El objetivo de la materia es la introducción de los aspectos básicos de la ingeniería de datos, fundamentalmente en el ámbito del Big Data. Las competencias adquiridas permitirán el análisis y la gestión eficiente de información heterogénea, tanto estructurada como no estructurada, dentro del desarrollo de aplicaciones IA, allí donde los métodos tradicionales muestren su insuficiencia. | | | | |

Competencias / Resultados del título

| Código | Competencias / Resultados del título |
|--------|---|
| A17 | CE16 - Conocimiento del proceso y las herramientas para el procesamiento y preparación de datos desde su adquisición o extracción, limpieza, transformación, carga, organización y acceso |
| B2 | CG02 - Abordar con éxito todas las etapas de un proyecto de Inteligencia Artificial |
| B3 | CG03 - Buscar y seleccionar la información útil necesaria para resolver problemas complejos, manejando con soltura las fuentes bibliográficas del campo |
| B4 | CG04 - Elaborar adecuadamente y con cierta originalidad composiciones escritas o argumentos motivados, redactar planes, proyectos de trabajo, artículos científicos y formular hipótesis razonables en el campo |
| B5 | CG05 - Trabajar en equipo, especialmente de carácter multidisciplinar, y ser hábiles en la gestión del tiempo, personas y toma de decisiones |
| B6 | CB01 - 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 |
| B7 | CB02 - 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 |
| B8 | CB03 - 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 |
| C3 | CT03 - 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 |
| C7 | CT07 - Desarrollar la capacidad de trabajar en equipos interdisciplinares o transdisciplinares, para ofrecer propuestas que contribuyan a un desarrollo sostenible ambiental, económico, político y social |
| C8 | CT08 - Valorar la importancia que tiene la investigación, la innovación y el desarrollo tecnológico en el avance socioeconómico y cultural de la sociedad |
| C9 | CT09 - Tener la capacidad de gestionar tiempos y recursos: desarrollar planes, priorizar actividades, identificar las críticas, establecer plazos y cumplirlos |

Resultados de aprendizaje

| Resultados de aprendizaje | Competencias / Resultados del título |
|---------------------------|--------------------------------------|
|---------------------------|--------------------------------------|



| | | | |
|---|------|--------------------------|-------------------|
| RA1 - Desarrollar la capacidad de analizar y modelar datos para su procesado en sistemas inteligentes | AM16 | BM6 BM7 | CM3 CM9 |
| RA2 - Conocer y comprender el proceso de extracción, limpieza, transformación, carga y preprocesado de datos. | AM16 | BM2 BM3 BM8 | CM3 CM7 CM9 |
| RA3 - Conocer y saber utilizar bases de datos multidimensionales y de tipo NoSQL | | BM3 BM4 BM7 | CM8 |
| RA4 - Conocer los fundamentos de data lakes y data warehouses | | BM2 BM5 BM7 BM8 | CM3 CM7 CM8 |

| Contenidos | |
|------------------------------------|---|
| Tema | Subtema |
| Conceptos y fundamentos | Conceptos y definiciones básicas, problemas de carga eficiente en escenarios Big Data, almacenamiento de datos masivos y acceso a los mismos. |
| Técnicas de limpieza y preparación | Técnicas más comunes. Definición de flujos de procesamiento. Métricas de calidad. |
| EStructuras avanzadas | Data warehouses y BD multidimensionales, Data lakes, Bases de Datos NoSQL. |

| Planificación | | | | |
|--------------------------|-----------------------------|---|------------------------|---------------|
| Metodologías / pruebas | Competencias / Resultados | Horas lectivas (presenciales y virtuales) | Horas trabajo autónomo | Horas totales |
| Sesión magistral | B4 B5 C3 C9 | 12 | 0 | 12 |
| Prácticas de laboratorio | A17 B2 B5 B7 C3 | 10 | 30 | 40 |
| Prueba mixta | A17 B2 B3 B6 B7 B8 C7 C8 | 3 | 20 | 23 |
| 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 | El profesorado presenta un tema al alumnado con el objetivo de facilitar un conjunto de información con alcance concreto. |
| Prácticas de laboratorio | El profesorado de la materia plantea al alumnado un problema o problemas de carácter práctico cuya resolución requiere la comprensión y aplicación de los contenidos teórico-prácticos incluidos en los contenidos de la materia. El alumnado puede trabajar la solución a los problemas planteados de forma individual o en grupos. |
| Prueba mixta | El examen abarca todos los temas del curso. Los alumnos deben desarrollar, relacionar, organizar y presentar los conocimientos que poseen sobre cada tema dado en una respuesta razonada y bien articulada. |

| Atención personalizada | |
|------------------------|-------------|
| Metodologías | Descripción |
| | |



| | |
|--------------------------|---|
| Sesión magistral | Lección magistral: se atenderán las dudas relacionadas con las metodologías y los casos de estudio analizados en las lecciones. |
| Prácticas de laboratorio | Prácticas de laboratorio: se atenderán las dudas relacionadas con los casos de estudio a analizar. |

| Evaluación | | | |
|--------------------------|---------------------------|--|--------------|
| Metodologías | Competencias / Resultados | Descripción | Calificación |
| Prueba mixta | A17 B2 B3 B6 B7 B8 C7 C8 | El examen abarca todos los temas del curso. Los alumnos deben desarrollar, relacionar, organizar y presentar los conocimientos que poseen sobre cada tema dado en una respuesta razonada y bien articulada. Resultados de aprendizaje evaluados: RA1, RA3, RA4 | 40 |
| Prácticas de laboratorio | A17 B2 B5 B7 C3 | Entrega de prácticas de laboratorio planteadas a lo largo del curso en las fechas estipuladas previamente. Resultados de aprendizaje evaluados: RA3, RA4 | 60 |

| Observaciones evaluación |
|---|
| <p>SISTEMA DE EVALUACIÓN CONTINUAPRUEBA 1: Prácticas de laboratorio % Calificación:60%. % Mínimo: Para la liberación de esta parte de la asignatura el estudiante deberá obtener una calificación igual o superior a5 puntos (sobre 10).PRUEBA 2: Examen de preguntas de desarrollo % Calificación: 40%. % Mínimo: Para la liberación de esta parte de la asignatura el estudiante deberá obtener una calificación igual o superior a5 puntos (sobre 10).SISTEMA DE EVALUACIÓN GLOBALProcedimiento para la elección de la modalidad de evaluación global: se considera que el estudiantado opta por el sistema de evaluación global si no se presenta a la Prueba 1 del sistema de evaluación continua.</p> <p>PRUEBA 1: Prácticas de laboratorio % Calificación:60%. % Mínimo: Para la liberación de esta parte de la asignatura el estudiante deberá obtener una calificación igual o superior a5 puntos (sobre 10).PRUEBA 2: Examen de preguntas de desarrollo % Calificación: 40%. % Mínimo: Para la liberación de esta parte de la asignatura el estudiante deberá obtener una calificación igual o superior a5 puntos (sobre 10).CRITERIOS DE EVALUACIÓN PARA CONVOCATORIA EXTRAORDINARIA (2ª OPORTUNIDAD) Y FIN DE CARRERASE emplearán los sistemas de evaluación continua y global expuestos anteriormente.PROCESO DE CALIFICACIÓN DE ACTASIndependientemente del sistema de evaluación y la convocatoria, en caso de no superar alguna parte de la evaluación, sila puntuación global fuese superior a 4 (sobre 10), la calificación en actas será 4.Todos los aspectos relacionados con ?dispensa académica?, ?dedicación al estudio?, ?permanencia? y ?fraude académico? se regirán de acuerdo con la normativa académica vigente de la UDC.</p> |

| Fuentes de información | |
|------------------------|--|
| Básica | <ul style="list-style-type: none"> - Sadalage, Fowler (2012). NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence. Addison-Wesley - Avi Silberschatz, Henry F. Korth, S. Sudarshan (2010). Database System Concepts. McGraw-Hill - Ihab F. Ilyas, Xu Chu, (2019). Data Cleaning. Association for Computing Machinery. ACM - Alex Gorelik (.). The Enterprise Big Data Lake: Delivering the Promise of Big Data and Data Science. O'Reilly |
| Complementaria | <ul style="list-style-type: none"> - Matt Casters, Roland Bouman, Jos van Dongen (2013). Pentaho Kettle Solutions: Building Open Source ETL Solutions with Pentaho Data Integration. Wiley |

| Recomendaciones |
|--|
| Asignaturas que se recomienda haber cursado previamente |
| |
| 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