

Gestión de Datos

Código: 106566
Créditos ECTS: 6

2025/2026

Titulación	Tipo	Curso
Inteligencia Artificial / Artificial Intelligence	OB	3

Contacto

Nombre: Javier Panadero Martínez

Correo electrónico: javier.panadero@uab.cat

Equipo docente

Haoyuan Li

Idiomas de los grupos

Puede consultar esta información al [final](#) del documento.

Prerrequisitos

Aunque no hay prerrequisitos formalmente establecidos y es responsabilidad de la propia asignatura proporcionar a los alumnos un medio para adquirir los conocimientos descritos en el apartado de contenidos de la asignatura. Es recomendable: un buen conocimiento de programación, del funcionamiento de un computador, del sistema operativo a nivel de usuario programador y de los sistemas de bases de datos.

Objetivos y contextualización

El objetivo de esta asignatura es conocer los conceptos fundamentales del diseño de las aplicaciones intensivas de datos respecto a su fiabilidad, escalabilidad y sostenibilidad.

Se estudiarán los sistemas distribuidos y los paradigmas de procesamiento masivo de datos. Se revisarán diversos modelos de programación para aplicaciones de procesamiento "batch", en memoria y "streaming". Se presentarán diversas arquitecturas de aplicaciones comunes y se presentarán conceptos relevantes como integridad, accesibilidad, fiabilidad, consistencia y seguridad en el procesamiento de datos a gran escala.

Competencias

- Analizar y resolver problemas de forma efectiva, generando propuestas innovadoras y creativas para alcanzar los objetivos.

- Conceptualizar y modelar alternativas de soluciones complejas a problemas de aplicación de la inteligencia artificial en diferentes ámbitos, y planificar y gestionar proyectos para el diseño y desarrollo de prototipos que demuestren la validez del sistema propuesto.
- Conocer y utilizar de forma eficiente las técnicas y herramientas de representación, manipulación, análisis y gestión de datos a gran escala.
- Introducir cambios en los métodos y los procesos del ámbito de conocimiento para dar respuestas innovadoras a las necesidades y demandas de la sociedad.
- Que los estudiantes hayan demostrado poseer y comprender conocimientos en un área de estudio que parte de la base de la educación secundaria general, y se suele encontrar a un nivel que, si bien se apoya en libros de texto avanzados, incluye también algunos aspectos que implican conocimientos procedentes de la vanguardia de su campo de estudio.
- Trabajar cooperativamente para la consecución de objetivos comunes, asumiendo la propia responsabilidad y respetando el rol de los diferentes miembros del equipo.

Resultados de aprendizaje

1. Analizar y resolver problemas de forma efectiva, generando propuestas innovadoras y creativas para alcanzar los objetivos.
2. Concebir, diseñar e implementar procesos de recopilación y anotación de datos adecuados al problema concreto a resolver.
3. Conocer los conceptos básicos de sistemas distribuidos de datos y el uso de herramientas de procesamiento masivo de datos.
4. Elegir los métodos de almacenamiento más apropiados que permitan la posterior recuperación y análisis de datos eficiente.
5. Proponer nuevos métodos o soluciones alternativas fundamentadas.
6. Que los estudiantes hayan demostrado poseer y comprender conocimientos en un área de estudio que parte de la base de la educación secundaria general, y se suele encontrar a un nivel que, si bien se apoya en libros de texto avanzados, incluye también algunos aspectos que implican conocimientos procedentes de la vanguardia de su campo de estudio.
7. Trabajar cooperativamente para la consecución de objetivos comunes, asumiendo la propia responsabilidad y respetando el rol de los diferentes miembros del equipo.

Contenido

1-Introducción a las aplicaciones masivas de datos

2-Conceptos fundamentales del tratamiento de datos en entornos masivos: fiabilidad, escalabilidad, sostenibilidad. Modelos de datos y lenguajes de consulta.

3-Gestión de grandes volúmenes de datos. Data warehousing. Principios de los sistemas Data Warehousing, Business intelligence, modelización multidimensional, operadores OLAP, esquema en estrella, proceso ETL

4-Introducción a las bases de datos en memoria con Redis

5-Aplicaciones con grandes volúmenes de datos con herramientas Apache. Introducción a las herramientas y ecosistema Apache Spark: Datasets, MLlib

Actividades formativas y Metodología

Título	Horas	ECTS	Resultados de aprendizaje
--------	-------	------	---------------------------

Tipo: Dirigidas

Estudio de casos de uso	9	0,36	1, 2, 5, 7
Prácticas	12	0,48	2, 3, 4, 7
Teoría	20	0,8	2, 3, 4, 6
Tipo: Autónomas			
Estudio autónomo	30	1,2	1, 3, 4, 6
Preparación de casos de estudio	20	0,8	1, 2, 3, 4, 5, 7
Preparación de laboratorios	32	1,28	1, 2, 7

En el desarrollo de la asignatura se podrán diferenciar los siguientes tipos de actividades docentes:

Clases teóricas: exposición de la parte teórica de cada tema del programa. La estructura típica de una clase magistral de este tipo será la siguiente: en primer lugar se hará una introducción donde se presentarán brevemente los objetivos de la exposición y los contenidos a tratar. A continuación, se desgranarán los contenidos objeto de estudio, incluyendo exposiciones narrativas, desarrollos formales que proporcionen los fundamentos teóricos, e intercalando ejemplos que ilustren la aplicación de los contenidos expuestos. Finalmente, el/la profesor/a expondrá las conclusiones de los contenidos. Durante todo el curso habrá evaluación continuada de grupos de temas.

Clases en el laboratorio: La parte práctica de los temas teóricos quedará completada con sesiones en el laboratorio, donde los grupos de estudiantes desarrollarán una serie de programas y deberán intentar resolver un problema concreto que recibirá al comenzar el temario. Algunos de estos ejercicios se deberán entregar a la clase en las fechas especificadas. Las prácticas se desarrollarán en grupos de dos estudiantes. Las clases incluyen varias sesiones en el laboratorio, de 2 horas de duración, donde el/la estudiante realizará el desarrollo de los ejercicios.

Problemas: revisión de casos prácticos. Al final de la asignatura se presentará una lista de casos prácticos con retos de datos relacionados con objetivos de análisis de negocio a realizar. Los estudiantes trabajarán en grupo para presentar una serie de conclusiones de su estudio de forma oral en las sesiones finales. Se realizará una evaluación en grupo de los trabajos presentados.

Este planteamiento del trabajo está orientado a promover un aprendizaje activo y desarrollar las competencias de capacidad de organización y planificación, comunicación oral y escrita, trabajo en equipo y razonamiento crítico. La calidad de los ejercicios realizados, de su presentación y de su funcionamiento se valorará especialmente.

Nota: se reservarán 15 minutos de una clase dentro del calendario establecido por el centro o por la titulación para que el alumnado rellene las encuestas de evaluación de la actuación del profesorado y de evaluación de la asignatura o módulo.

Evaluación

Actividades de evaluación continuada

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Control individual 1	30%	2	0,08	1, 3, 6
Control individual 2	30%	2	0,08	1, 3, 6

Laboratorios de prácticas	40%	23	0,92	1, 2, 3, 4, 5, 7
---------------------------	-----	----	------	------------------

Esta asignatura no contempla el sistema de evaluación única. Los estudiantes repetidores deberán realizar todas las actividades planificadas, tanto teóricas como prácticas, es decir, no habrá un tratamiento diferenciado para los estudiantes repetidores.

La asignatura consta de tres partes: Teoría, Problemas y Prácticas. La parte de Teoría y Problemas representa el 60% de la nota final, mientras que la parte de Prácticas representa el 40% restante.

Las fechas de las pruebas de evaluación continua y prácticas se publicarán en el campus virtual al principio de curso y pueden estar sujetas a posibles cambios de programación por motivos de adaptación a posibles incidencias. Siempre se informará en el campus virtual sobre dichos cambios, ya que se entiende que el CV es el mecanismo habitual de intercambio de información entre profesorado y estudiantes.

Se concederán las matrículas de honor resultantes de calcular el cinco por ciento o fracción de los alumnos matriculados en todos los grupos docentes de la asignatura. Solo podrán otorgarse a estudiantes que hayan obtenido una calificación final igual o superior a 9.

La forma en que se evaluará cada una de las distintas partes de que consta la asignatura (Teoría, Problemas y Prácticas) se detalla a continuación:

Teoría y Problemas

La asignatura seguirá una metodología de evaluación continua que permitirá ir descartando temario a medida que el estudiante avance en el curso. Está prevista la realización de dos pruebas escritas de evaluación continua:

- La primera prueba (P1) se realizará durante la semana de parciales.
- La segunda prueba (P2) se realizará durante la semana de exámenes finales.

La fecha exacta de las pruebas se publicará al inicio del curso y podrá variar en función de posibles incidencias que puedan surgir durante el semestre. Siempre se informará en el campus virtual sobre estos cambios, ya que se entiende que el CV es el mecanismo habitual de intercambio de información entre profesorado y estudiantes.

Cada prueba tendrá un peso del 30% respecto a la nota final de la asignatura.

Para poder realizar la segunda prueba de evaluación continua (P2), será necesario que el estudiante obtenga una nota mínima de 3,5 en la primera prueba (P1). En caso contrario, deberá realizar el examen de recuperación (ER), que incluirá preguntas de todo el contenido del curso. Además, si la media entre las dos pruebas parciales (P1 y P2) es inferior a 5, el estudiante también deberá realizar el examen de recuperación para optar a aprobar la asignatura.

Para cada prueba, se indicará lugar, fecha y hora de revisión, en la que el estudiante podrá revisar la actividad con el profesor. Si el estudiante no se presenta a esta revisión, no se revisará posteriormente dicha actividad.

El estudiante que quiera asistir a la revisión deberá avisar, como mínimo, con 24 horas de antelación respecto a la fecha fijada, mediante un correo electrónico dirigido a su profesor de teoría. En caso de no avisar dentro de este plazo, no se revisará su examen.

Durante la sesión de revisión no se explicarán ni se resolverán los ejercicios del examen. Solo se mostrará el examen para que el estudiante pueda identificar en qué partes ha fallado y entender el motivo de su calificación.

Las soluciones de los exámenes no se publicarán en el Campus Virtual. Si un estudiante quiere conocer la solución de algún ejercicio, deberá solicitar una tutoría una vez finalizado el proceso de revisión.

Examen de recuperación

El examen de recuperación solo podrán realizarlo los estudiantes que no hayan superado la evaluación continua, ya sea porque no han alcanzado la nota mínima de 5 puntos sobre 10, o porque no la han seguido.

Este examen incluirá preguntas de todo el temario del curso y tendrá una puntuación máxima de 7 puntos. La nota mínima para hacer media con la parte práctica es un 5. Obtener una nota inferior a 5 supondrá suspender la asignatura.

Si se detecta cualquier intento de copia en alguna prueba evaluable, tanto durante la realización como en el proceso de corrección, la nota final será de 3 en el expediente, y se procederá a abrir un expediente disciplinario, que quedará reflejado en el registro académico del estudiante.

El equipo docente se reserva el derecho de modificar el formato de las pruebas parciales y finales según considere oportuno, sin que estas tengan que ser similares a las de años anteriores.

Prácticas de laboratorio

La evaluación de las prácticas se realizará a partir del desarrollo del trabajo hecho durante las sesiones de laboratorio y de las memorias redactadas para cada una de las sesiones. Las prácticas se desarrollarán en grupos de 3 estudiantes obligatoriamente.

La asistencia a las sesiones prácticas es obligatoria. No asistir a alguna sesión supondrá suspender la parte práctica y, en consecuencia, la asignatura. En caso de incidencia justificada, deberá comunicarse previamente al profesor responsable y aportar un justificante oficial debidamente firmado dentro del plazo establecido. Siempre deberá avisarse antes de la sesión.

Debe aclararse que los viajes personales, así como los motivos laborales, no se consideran incidencias justificables, ya que el estudiante dispone del calendario de prácticas desde el inicio del curso para evitar este tipo de situaciones.

Las incidencias puntuales justificables, que no sean por enfermedad, deberán recuperarse en otro turno durante la misma semana. Solo estarán exentos de recuperación los estudiantes que justifiquen una enfermedad. En cualquier caso, el hecho de no asistir al turno de prácticas asignado y, por tanto, impedir el trabajo conjunto con su grupo, supondrá que el estudiante deba realizar la práctica de manera individual.

Es obligatorio estar presente durante toda la sesión de prácticas. Se pasará lista al inicio de la sesión y, antes de marcharse, el profesor preguntará a los estudiantes por el trabajo realizado para valorar el trabajo hecho durante la sesión, momento en que se volverá a pasar lista. El profesor preguntará a cada miembro del grupo para establecer la nota individual de laboratorio de la sesión. El porcentaje de cada parte de la práctica se detallará en la normativa específica de la asignatura.

Todas las prácticas tienen la misma puntuación. Las puntuaciones específicas de cada práctica se detallarán en el enunciado correspondiente. Es responsabilidad del estudiante leerlo detenidamente para conocer las puntuaciones asignadas a cada apartado. Asimismo, también es responsabilidad del estudiante apuntarse en el listado de asistencia de cada sesión.

Un retraso superior a 15 minutos desde el inicio de la práctica supondrá un "no presentado" en la sesión, sin posibilidad de recuperación. Esta condición no se aplicará a los estudiantes que puedan acreditar el retraso con un justificante oficial (por ejemplo, un certificado de asistencia médica).

Las prácticas no son recuperables. Para aprobar esta parte de la asignatura, es necesario obtener una nota media mínima de 5. No se establece ninguna nota mínima para las prácticas individuales para poder calcular la media global.

Plagio y copias

Sin perjuicio de otras medidas disciplinarias que se consideren oportunas, y de acuerdo con la normativa académica vigente, las irregularidades cometidas por un estudiante que puedan conducir a una variación de la calificación en una actividad evaluable se calificarán con un cero (0). Las actividades de evaluación calificadas

de esta forma y por este procedimiento no serán recuperables. Si es necesario superar cualquiera de estas actividades de evaluación para aprobar la asignatura, esta asignatura quedará suspendida directamente, sin oportunidad de recuperarla en el mismo curso. Estas irregularidades incluyen, entre otras:

- la copia total o parcial de una práctica, informe o cualquier otra actividad de evaluación; dejar copiar;
- el uso no autorizado de IA (por ejemplo, Copilot, ChatGPT o equivalentes), etc. en cualquiera de las actividades de evaluación implicará suspenderla con un cero;
- presentar un trabajo en grupo no realizado íntegramente por los miembros del grupo (aplicable a todos los miembros, no solo a los que no han trabajado);
- presentar como propios materiales elaborados por un tercero, aunque sean traducciones o adaptaciones, y en general trabajos con elementos no originales y exclusivos del estudiante;
- tener dispositivos de comunicación (como teléfonos móviles, smartwatches, bolígrafos con cámara, etc.) accesibles durante las pruebas de evaluación teórico-prácticas individuales (exámenes);
- hablar con compañeros durante las pruebas de evaluación teórico-prácticas individuales (exámenes);
- copiar o intentar copiar a otros alumnos durante las pruebas de evaluación teórico-prácticas (exámenes);
- utilizar o intentar utilizar escritos relacionados con la materia durante la realización de las pruebas de evaluación teórico-prácticas (exámenes), cuando estos no hayan sido explícitamente permitidos.

En resumen: copiar, dejar copiar o plagiar (o intentarlo) en cualquiera de las actividades de evaluación equivale a un SUSPENSO, no compensable y sin convalidaciones de partes de la asignatura en cursos posteriores.

En caso de no superar la asignatura por el hecho de que alguna de las actividades de evaluación no alcanza la nota mínima requerida, la nota numérica del expediente será el valor menor entre 4,5 y la media ponderada de las notas. Con las excepciones de que se otorgará la calificación de "No Evaluado" a los estudiantes que no participen en ninguna de las actividades de evaluación, y que la nota numérica del expediente será el valor menor entre 3,0 y la media ponderada de las notas en caso de que el estudiante haya cometido irregularidades en un acto de evaluación (y por tanto no será posible aprobar por compensación).

Bibliografía

Designing Data intensive applications - Martin Kleppmann, O'Reilly, 2017

The Data warehouse ETL toolkit - Ralph Kimball, Joe Caserta. Wiley, 2004

Spark, the definitive guide, Big data processing made simple. Bill Chambers and Matei Zaharia, O'Reilly, 2018

Learning Spark - Lightning fast data analysis - Holden Karau, Andi Konwinski, Patrick Wendell, Matei Zaharia, O'Reilly, 2015

Beginning Scala - Layka, Vishal. Apress; 2nd ed. 2015.

Redis in Action - Josiah L. Carlson. Manning, 2013.

Software

Se utilizarán los servicios en la nube proporcionados por la Escuela de Ingeniería en la plataforma OpenNebula

Grupos e idiomas de la asignatura

La información proporcionada es provisional hasta el 30 de noviembre de 2025. A partir de esta fecha, podrá consultar el idioma de cada grupo a través de este [enlace](#). Para acceder a la información, será necesario introducir el CÓDIGO de la asignatura

Nombre	Grupo	Idioma	Semestre	Turno
(PAUL) Prácticas de aula	711	Inglés	primer cuatrimestre	tarde
(PLAB) Prácticas de laboratorio	711	Inglés	primer cuatrimestre	tarde
(TE) Teoría	71	Inglés	primer cuatrimestre	tarde