

## Guía docente

### 270220 - CAI - Búsqueda y Análisis de la Información

Última modificación: 07/07/2026

**Unidad responsable:** Facultad de Informática de Barcelona  
**Unidad que imparte:** 723 - CS - Departamento de Ciencias de la Computación.  
**Titulación:** GRADO EN CIENCIA E INGENIERÍA DE DATOS (Plan 2017). (Asignatura obligatoria).  
**Curso:** 2026      **Créditos ECTS:** 6.0      **Idiomas:** Inglés

#### PROFESORADO

---

**Profesorado responsable:** MARTA ARIAS VICENTE  
**Otros:** Primer quadrimestre:  
MARTA ARIAS VICENTE - 11, 12  
DAVID GARCIA SORIANO - 11, 12

#### CAPACIDADES PREVIAS

---

Las dadas por las asignaturas de los Cuatrimestres 1 a 4 del grado. La asignatura es en buena parte "comprensiva" con mucho el hecho anteriormente, y especialmente conceptos de matemática (discreta, álgebra, algo de cálculo), probabilidad y estadística, algoritmia general (en particular, de grafos), aprendizaje automático y análisis de datos, bases de datos, y sistemas de computadores distribuidos y paralelos. Los laboratorios se realizan en el lenguaje de programación python.

#### COMPETENCIAS DE LA TITULACIÓN A LAS QUE CONTRIBUYE LA ASIGNATURA

---

##### Específicas:

CE1. Utilizar con destreza los conceptos y métodos matemáticos que subyacen los problemas de la ciencia y la ingeniería de los datos.  
CE4. Utilizar los sistemas de computación actuales, incluidos sistemas de alto rendimiento, para el proceso de grandes volúmenes de datos desde el conocimiento de su estructura, funcionamiento y particularidades.  
CE6. Construir o utilizar sistemas de procesamiento y comprensión del lenguaje escrito, integrándolo en otros sistemas dirigidos por los datos. Diseñar sistemas de búsqueda de información textual o hipertextual y de análisis de redes sociales.  
CE7. Demostrar conocimiento y capacidad de aplicación de las herramientas necesarias para el almacenaje, el procesamiento y el acceso a los datos.

##### Genéricas:

CG2. Elegir y aplicar los métodos y técnicas más adecuados a un problema definido por datos que representen un reto por su volumen, velocidad, variedad o heterogeneidad, incluidos métodos informáticos, matemáticos, estadísticos y de procesamiento de la señal.  
CG3. Trabajar en equipos y proyectos multidisciplinares relacionados con el procesamiento y explotación de datos complejos, interactuando fluidamente con ingenieros y profesionales de otras disciplinas.  
CG4. Identificar oportunidades para aplicaciones innovadoras orientadas a datos en entornos tecnológicos en continua evolución.  
CG5. Poder recurrir a conocimientos fundamentales y metodologías de trabajo sólidas adquiridos durante los estudios para adaptarse a los nuevos escenarios tecnológicos del futuro.

##### Transversales:

CT4. Trabajo en equipo. Ser capaz de trabajar como miembro de un equipo interdisciplinar, ya sea como un miembro más o realizando tareas de dirección, con la finalidad de contribuir a desarrollar proyectos con pragmatismo y sentido de la responsabilidad, asumiendo compromisos teniendo en cuenta los recursos disponibles.  
CT5. Uso solvente de los recursos de información. Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de especialidad y valorar de forma crítica los resultados de dicha gestión.  
CT6. Aprendizaje autónomo. Detectar deficiencias en el propio conocimiento y superarlas mediante la reflexión crítica y la elección de la mejor actuación para ampliar dicho conocimiento.  
CT7. Tercera lengua. Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

#### **Básicas:**

CB2. Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio.

CB3. Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética.

CB4. Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.

## **METODOLOGÍAS DOCENTES**

---

Classes "de teoria" expositives per part del professor. Es proposaran un cert nombre d'exercicis a resoldre fora de classe per a la propera sessió.

Classes "de teoria" dedicades a la resolució. Es comentaran en comú les solucions dels exercicis proposats a la/les sessions precedents. S'esperarà que els estudiants hagin intentat resoldre'ls.

Classes "de laboratori": A partir d'un guió que rebran al principi de la sessió, els estudiants duran a terme alguna tasca amb ordinador per consolidar els conceptes vistos a les classes de "teoria". Típicament serà l'implementació i experimentació amb algun algorisme, o l'anàlisi d'algun conjunt de dades.

## **OBJETIVOS DE APRENDIZAJE DE LA ASIGNATURA**

---

1.Describir diferentes modelos para evaluar similaridad entre textos, y cómo se aplican a la búsqueda textual. Decidir cuál de los modelos es el más adecuado a un escenario concreto que involucra búsqueda de textos. Implementar los modelos desde cero (en un sistema muy básico) o sobre un sistema de indexación de textos que sea altamente escalable.

2.Describir las ventajas, de cara a efectuar búsquedas efectivas, de utilizar la información dada por los enlaces en estructuras hiperenllaçades, como la web, las redes sociales digitales, y la web semántica. Describir los principales parámetros que se usan para caracterizar estas estructuras enlazadas. Reproducir los algoritmos más utilizados para establecer importancia en estas estructuras (p.ej. pagerank), para descubrir estructura en ellas (por ejemplo, descubrimiento de comunidades) y para mejorar los resultados de búsqueda propuestas por un usuario. Implementar estos algoritmos desde cero en un sistema muy básico, o encima de sistemas de procesamiento de datos masivos de manera que puedan escalar.

3.Evaluar la efectividad de sistemas de búsqueda en sistemas complejos, describiéndola en términos de medidas duras como "recall" y "precisión" pero también en términos de medidas blandas como la satisfacción de los usuarios, la novedad y la finalización de tareas . Adaptar el funcionamiento y la presentación de sistemas de búsqueda de información con el feedback de la experiencia de usuarios metódicamente recogida.

4.Definir el problema de la recomendación y las diferencias con otros problemas relacionados con información previamente almacenada (búsqueda, aprendizaje, ...). Describir los principales abordajes al problema de la recomendación de items y las ventajas e inconvenientes de cada uno. Describir los principales algoritmos de cada uno de los abordajes. Ser capaz de implementar versiones básicas desde cero, o versiones avanzadas encima de sistemas de procesamiento masivos de datos. Evaluar la efectividad de sistemas de recomendación, tanto en términos de medidas duras como de medidas blandas como la satisfacción del usuario. Decidir las formas de recomendación más adecuadas a escenarios reales simples, incluyendo la caracterización de los usuarios potenciales.

5.Recurrir de manera a paradigmas algorítmicos conocidos para tratar con problemas de datos caracterizados por alto volumen y alta velocidad. Incluyen: algoritmos de "streaming" que tratan flujos de datos con poco tiempo por elemento, y poca memoria. Algoritmos para responder preguntas de proximidad, en particular con información geolocalizada. Algoritmos que utilizan muestreo para extraer conclusiones fiables sobre grandes volúmenes de datos. Integración de las técnicas vistas en el resto del curso con técnicas algorítmicas de otras asignaturas, tales como "machine learning", "clustering" y "pattern mining". Técnicas para tratar con datos sensibles, tales como anonimización y "privacy-preserving machine learning". "Caching" consistente y distribuido.

6.Integrar las técnicas descritas en los anteriores objetivos en un proyecto pequeño pero realista. Tener la capacidad para diseñar la arquitectura de un sistema complejo y elegir las técnicas, y las tecnologías vistas con anterioridad durante el curso que se aplicarían. El objetivo no es finalizar la implementación del sistema, sino llegar a un nivel de detalle del diseño que permitiera encargó a un equipo de programación su finalización.

7.Evaluar de manera elemental las implicaciones de los sistemas que se aprenden a construir en la asignatura en cuanto a privacidad, seguridad, ética y derechos de las personas. Se entiende por "de manera elemental" ser capaz de detectar que estas implicaciones son suficientemente significativas como para buscar la opinión de un expertos en la materia, en particular en relación al RGPD y la necesidad de llevar a cabo análisis de riesgo y impacto.

## HORAS TOTALES DE DEDICACIÓN DEL ESTUDIANTE

| Tipo                       | Horas | Porcentaje |
|----------------------------|-------|------------|
| Horas grupo pequeño        | 30,0  | 20.00      |
| Horas grupo grande         | 30,0  | 20.00      |
| Horas aprendizaje autónomo | 90,0  | 60.00      |

**Dedicación total:** 150 h

## CONTENIDOS

### Búsqueda y análisis de información textual

**Descripción:**

Modelos booleano y vectorial. Búsqueda basada en palabras clave. Preproceso de los textos. Indexación. Evaluación de estrategias de búsqueda. Técnicas eficientes de búsqueda basadas en funciones de hash (locality sensitive hashing). Modelos de semántica latente para la representación de textos (LSA).

### Búsqueda y análisis en estructuras enlazadas

**Descripción:**

La web: Algorismes d'avaluació en estructures hiperenllaçades. "Crawling" i "scraping". Xarxes socials: Mesures de centralitat. Comunitats. Influència. Web semàntica.

### Algoritmos para datos masivos

**Descripción:**

Resums (sketches) i fluxos de dades (streaming). Mostatge (sampling). Preguntes de proximitat. Dades geolocalitzades. "Caching" consistent i distribuït. Tractament de dades sensibles: anonimització, "end-to-end encryption" i "privacy-preserving machine learning"

## ACTIVIDADES

### Activitat sobre el contingut "Cerca i anàlisi d'informació textual"

#### Objetivos específicos:

1, 3, 6, 7

#### Competencias relacionadas:

CG2. Elegir y aplicar los métodos y técnicas más adecuados a un problema definido por datos que representen un reto por su volumen, velocidad, variedad o heterogeneidad, incluidos métodos informáticos, matemáticos, estadísticos y de procesado de la señal.

CG4. Identificar oportunidades para aplicaciones innovadoras orientadas a datos en entornos tecnológicos en continua evolución.

CG5. Poder recurrir a conocimientos fundamentales y metodologías de trabajo sólidas adquiridos durante los estudios para adaptarse a los nuevos escenarios tecnológicos del futuro.

CG3. Trabajar en equipos y proyectos multidisciplinares relacionados con el procesado y explotación de datos complejos, interactuando fluidamente con ingenieros y profesionales de otras disciplinas.

CE1. Utilizar con destreza los conceptos y métodos matemáticos que subyacen los problemas de la ciencia y la ingeniería de los datos.

CE6. Construir o utilizar sistemas de procesado y comprensión del lenguaje escrito, integrándolo en otros sistemas dirigidos por los datos. Diseñar sistemas de búsqueda de información textual o hipertextual y de análisis de redes sociales.

CE7. Demostrar conocimiento y capacidad de aplicación de las herramientas necesarias para el almacenaje, el procesamiento y el acceso a los datos.

CE4. Utilizar los sistemas de computación actuales, incluidos sistemas de alto rendimiento, para el proceso de grandes volúmenes de datos desde el conocimiento de su estructura, funcionamiento y particularidades.

CT5. Uso solvente de los recursos de información. Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de especialidad y valorar de forma crítica los resultados de dicha gestión.

CT6. Aprendizaje autónomo. Detectar deficiencias en el propio conocimiento y superarlas mediante la reflexión crítica y la elección de la mejor actuación para ampliar dicho conocimiento.

CT4. Trabajo en equipo. Ser capaz de trabajar como miembro de un equipo interdisciplinar, ya sea como un miembro más o realizando tareas de dirección, con la finalidad de contribuir a desarrollar proyectos con pragmatismo y sentido de la responsabilidad, asumiendo compromisos teniendo en cuenta los recursos disponibles.

CT7. Tercera lengua. Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

CB3. Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética.

CB4. Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.

CB2. Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio.

#### Dedicación: 30h

Grupo grande/Teoría: 8h

Grupo pequeño/Laboratorio: 10h

Aprendizaje autónomo: 12h

### Activitat sobre el contingut "Cerca i anàlisi en estructures enllaçades"

#### Objetivos específicos:

2, 6, 7

#### Competencias relacionadas:

CG2. Elegir y aplicar los métodos y técnicas más adecuados a un problema definido por datos que representen un reto por su volumen, velocidad, variedad o heterogeneidad, incluidos métodos informáticos, matemáticos, estadísticos y de procesado de la señal.

CG4. Identificar oportunidades para aplicaciones innovadoras orientadas a datos en entornos tecnológicos en continua evolución.

CG5. Poder recurrir a conocimientos fundamentales y metodologías de trabajo sólidas adquiridos durante los estudios para adaptarse a los nuevos escenarios tecnológicos del futuro.

CG3. Trabajar en equipos y proyectos multidisciplinares relacionados con el procesado y explotación de datos complejos, interactuando fluidamente con ingenieros y profesionales de otras disciplinas.

CE1. Utilizar con destreza los conceptos y métodos matemáticos que subyacen los problemas de la ciencia y la ingeniería de los datos.

CE6. Construir o utilizar sistemas de procesado y comprensión del lenguaje escrito, integrándolo en otros sistemas dirigidos por los datos. Diseñar sistemas de búsqueda de información textual o hipertextual y de análisis de redes sociales.

CE7. Demostrar conocimiento y capacidad de aplicación de las herramientas necesarias para el almacenaje, el procesamiento y el acceso a los datos.

CE4. Utilizar los sistemas de computación actuales, incluidos sistemas de alto rendimiento, para el proceso de grandes volúmenes de datos desde el conocimiento de su estructura, funcionamiento y particularidades.

CT5. Uso solvente de los recursos de información. Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de especialidad y valorar de forma crítica los resultados de dicha gestión.

CT6. Aprendizaje autónomo. Detectar deficiencias en el propio conocimiento y superarlas mediante la reflexión crítica y la elección de la mejor actuación para ampliar dicho conocimiento.

CT4. Trabajo en equipo. Ser capaz de trabajar como miembro de un equipo interdisciplinar, ya sea como un miembro más o realizando tareas de dirección, con la finalidad de contribuir a desarrollar proyectos con pragmatismo y sentido de la responsabilidad, asumiendo compromisos teniendo en cuenta los recursos disponibles.

CT7. Tercera lengua. Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

CB3. Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética.

CB4. Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.

CB2. Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio.

#### Dedicación: 28h

Grupo grande/Teoría: 8h

Grupo pequeño/Laboratorio: 8h

Aprendizaje autónomo: 12h

## Activitat sobre el contingut "Algorismes per a dades massives"

### Objetivos específicos:

5, 6, 7

### Competencias relacionadas:

CG2. Elegir y aplicar los métodos y técnicas más adecuados a un problema definido por datos que representen un reto por su volumen, velocidad, variedad o heterogeneidad, incluidos métodos informáticos, matemáticos, estadísticos y de procesado de la señal.

CG4. Identificar oportunidades para aplicaciones innovadoras orientadas a datos en entornos tecnológicos en continua evolución.

CG5. Poder recurrir a conocimientos fundamentales y metodologías de trabajo sólidas adquiridos durante los estudios para adaptarse a los nuevos escenarios tecnológicos del futuro.

CG3. Trabajar en equipos y proyectos multidisciplinares relacionados con el procesado y explotación de datos complejos, interactuando fluidamente con ingenieros y profesionales de otras disciplinas.

CE1. Utilizar con destreza los conceptos y métodos matemáticos que subyacen los problemas de la ciencia y la ingeniería de los datos.

CE6. Construir o utilizar sistemas de procesado y comprensión del lenguaje escrito, integrándolo en otros sistemas dirigidos por los datos. Diseñar sistemas de búsqueda de información textual o hipertextual y de análisis de redes sociales.

CE7. Demostrar conocimiento y capacidad de aplicación de las herramientas necesarias para el almacenaje, el procesamiento y el acceso a los datos.

CE4. Utilizar los sistemas de computación actuales, incluidos sistemas de alto rendimiento, para el proceso de grandes volúmenes de datos desde el conocimiento de su estructura, funcionamiento y particularidades.

CT5. Uso solvente de los recursos de información. Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de especialidad y valorar de forma crítica los resultados de dicha gestión.

CT6. Aprendizaje autónomo. Detectar deficiencias en el propio conocimiento y superarlas mediante la reflexión crítica y la elección de la mejor actuación para ampliar dicho conocimiento.

CT4. Trabajo en equipo. Ser capaz de trabajar como miembro de un equipo interdisciplinar, ya sea como un miembro más o realizando tareas de dirección, con la finalidad de contribuir a desarrollar proyectos con pragmatismo y sentido de la responsabilidad, asumiendo compromisos teniendo en cuenta los recursos disponibles.

CT7. Tercera lengua. Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

CB3. Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética.

CB4. Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.

CB2. Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio.

### Dedicación: 28h

Grupo grande/Teoría: 8h

Grupo pequeño/Laboratorio: 8h

Aprendizaje autónomo: 12h

## Examen parcial

### Descripción:

Examen a resolver presencialmente durante la semana de parciales.

### Objetivos específicos:

1, 2, 3

### Competencias relacionadas:

CG2. Elegir y aplicar los métodos y técnicas más adecuados a un problema definido por datos que representen un reto por su volumen, velocidad, variedad o heterogeneidad, incluidos métodos informáticos, matemáticos, estadísticos y de procesado de la señal.

CG4. Identificar oportunidades para aplicaciones innovadoras orientadas a datos en entornos tecnológicos en continua evolución.

CG5. Poder recurrir a conocimientos fundamentales y metodologías de trabajo sólidas adquiridos durante los estudios para adaptarse a los nuevos escenarios tecnológicos del futuro.

CG3. Trabajar en equipos y proyectos multidisciplinares relacionados con el procesado y explotación de datos complejos, interactuando fluidamente con ingenieros y profesionales de otras disciplinas.

CE1. Utilizar con destreza los conceptos y métodos matemáticos que subyacen los problemas de la ciencia y la ingeniería de los datos.

CE6. Construir o utilizar sistemas de procesado y comprensión del lenguaje escrito, integrándolo en otros sistemas dirigidos por los datos. Diseñar sistemas de búsqueda de información textual o hipertextual y de análisis de redes sociales.

CE7. Demostrar conocimiento y capacidad de aplicación de las herramientas necesarias para el almacenaje, el procesamiento y el acceso a los datos.

CE4. Utilizar los sistemas de computación actuales, incluidos sistemas de alto rendimiento, para el proceso de grandes volúmenes de datos desde el conocimiento de su estructura, funcionamiento y particularidades.

CT5. Uso solvente de los recursos de información. Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de especialidad y valorar de forma crítica los resultados de dicha gestión.

CT6. Aprendizaje autónomo. Detectar deficiencias en el propio conocimiento y superarlas mediante la reflexión crítica y la elección de la mejor actuación para ampliar dicho conocimiento.

CT4. Trabajo en equipo. Ser capaz de trabajar como miembro de un equipo interdisciplinar, ya sea como un miembro más o realizando tareas de dirección, con la finalidad de contribuir a desarrollar proyectos con pragmatismo y sentido de la responsabilidad, asumiendo compromisos teniendo en cuenta los recursos disponibles.

CT7. Tercera lengua. Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

CB3. Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética.

CB4. Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.

CB2. Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio.

### Dedicación: 15h

Actividades dirigidas: 3h

Aprendizaje autónomo: 12h

## Examen final

### Descripción:

Resolución del examen final, una vez terminado el período de docencia, en un aula y un tiempo determinado.

### Objetivos específicos:

1, 2, 3, 4, 5, 7

### Competencias relacionadas:

CG2. Elegir y aplicar los métodos y técnicas más adecuados a un problema definido por datos que representen un reto por su volumen, velocidad, variedad o heterogeneidad, incluidos métodos informáticos, matemáticos, estadísticos y de procesado de la señal.

CG4. Identificar oportunidades para aplicaciones innovadoras orientadas a datos en entornos tecnológicos en continua evolución.

CG5. Poder recurrir a conocimientos fundamentales y metodologías de trabajo sólidas adquiridos durante los estudios para adaptarse a los nuevos escenarios tecnológicos del futuro.

CG3. Trabajar en equipos y proyectos multidisciplinares relacionados con el procesado y explotación de datos complejos, interactuando fluidamente con ingenieros y profesionales de otras disciplinas.

CE1. Utilizar con destreza los conceptos y métodos matemáticos que subyacen los problemas de la ciencia y la ingeniería de los datos.

CE6. Construir o utilizar sistemas de procesado y comprensión del lenguaje escrito, integrándolo en otros sistemas dirigidos por los datos. Diseñar sistemas de búsqueda de información textual o hipertextual y de análisis de redes sociales.

CE7. Demostrar conocimiento y capacidad de aplicación de las herramientas necesarias para el almacenaje, el procesamiento y el acceso a los datos.

CE4. Utilizar los sistemas de computación actuales, incluidos sistemas de alto rendimiento, para el proceso de grandes volúmenes de datos desde el conocimiento de su estructura, funcionamiento y particularidades.

CT5. Uso solvente de los recursos de información. Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de especialidad y valorar de forma crítica los resultados de dicha gestión.

CT6. Aprendizaje autónomo. Detectar deficiencias en el propio conocimiento y superarlas mediante la reflexión crítica y la elección de la mejor actuación para ampliar dicho conocimiento.

CT4. Trabajo en equipo. Ser capaz de trabajar como miembro de un equipo interdisciplinar, ya sea como un miembro más o realizando tareas de dirección, con la finalidad de contribuir a desarrollar proyectos con pragmatismo y sentido de la responsabilidad, asumiendo compromisos teniendo en cuenta los recursos disponibles.

CT7. Tercera lengua. Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

CB3. Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética.

CB4. Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.

CB2. Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio.

### Dedicación: 17h

Actividades dirigidas: 3h

Aprendizaje autónomo: 14h

## Primer control de laboratorio

### Descripción:

Control de laboratorio que evalúa el contenido de las sesiones de laboratorio precedentes al control

### Objetivos específicos:

6

### Competencias relacionadas:

CG2. Elegir y aplicar los métodos y técnicas más adecuados a un problema definido por datos que representen un reto por su volumen, velocidad, variedad o heterogeneidad, incluidos métodos informáticos, matemáticos, estadísticos y de procesado de la señal.

CG4. Identificar oportunidades para aplicaciones innovadoras orientadas a datos en entornos tecnológicos en continua evolución.

CG5. Poder recurrir a conocimientos fundamentales y metodologías de trabajo sólidas adquiridos durante los estudios para adaptarse a los nuevos escenarios tecnológicos del futuro.

CG3. Trabajar en equipos y proyectos multidisciplinares relacionados con el procesado y explotación de datos complejos, interactuando fluidamente con ingenieros y profesionales de otras disciplinas.

CE1. Utilizar con destreza los conceptos y métodos matemáticos que subyacen los problemas de la ciencia y la ingeniería de los datos.

CE6. Construir o utilizar sistemas de procesado y comprensión del lenguaje escrito, integrándolo en otros sistemas dirigidos por los datos. Diseñar sistemas de búsqueda de información textual o hipertextual y de análisis de redes sociales.

CE7. Demostrar conocimiento y capacidad de aplicación de las herramientas necesarias para el almacenaje, el procesamiento y el acceso a los datos.

CE4. Utilizar los sistemas de computación actuales, incluidos sistemas de alto rendimiento, para el proceso de grandes volúmenes de datos desde el conocimiento de su estructura, funcionamiento y particularidades.

CT5. Uso solvente de los recursos de información. Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de especialidad y valorar de forma crítica los resultados de dicha gestión.

CT4. Trabajo en equipo. Ser capaz de trabajar como miembro de un equipo interdisciplinar, ya sea como un miembro más o realizando tareas de dirección, con la finalidad de contribuir a desarrollar proyectos con pragmatismo y sentido de la responsabilidad, asumiendo compromisos teniendo en cuenta los recursos disponibles.

CT7. Tercera lengua. Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

CB4. Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.

CB2. Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio.

### Dedicación: 16h

Actividades dirigidas: 2h

Aprendizaje autónomo: 14h

## Segundo control de laboratorio

### Descripción:

Evalúa el trabajo realizado durante las sesiones de laboratorio precedentes al control

### Objetivos específicos:

6

### Competencias relacionadas:

CG2. Elegir y aplicar los métodos y técnicas más adecuados a un problema definido por datos que representen un reto por su volumen, velocidad, variedad o heterogeneidad, incluidos métodos informáticos, matemáticos, estadísticos y de procesado de la señal.

CG4. Identificar oportunidades para aplicaciones innovadoras orientadas a datos en entornos tecnológicos en continua evolución.

CG5. Poder recurrir a conocimientos fundamentales y metodologías de trabajo sólidas adquiridos durante los estudios para adaptarse a los nuevos escenarios tecnológicos del futuro.

CG3. Trabajar en equipos y proyectos multidisciplinares relacionados con el procesado y explotación de datos complejos, interactuando fluidamente con ingenieros y profesionales de otras disciplinas.

CE1. Utilizar con destreza los conceptos y métodos matemáticos que subyacen los problemas de la ciencia y la ingeniería de los datos.

CE6. Construir o utilizar sistemas de procesado y comprensión del lenguaje escrito, integrándolo en otros sistemas dirigidos por los datos. Diseñar sistemas de búsqueda de información textual o hipertextual y de análisis de redes sociales.

CE7. Demostrar conocimiento y capacidad de aplicación de las herramientas necesarias para el almacenaje, el procesamiento y el acceso a los datos.

CE4. Utilizar los sistemas de computación actuales, incluidos sistemas de alto rendimiento, para el proceso de grandes volúmenes de datos desde el conocimiento de su estructura, funcionamiento y particularidades.

CT5. Uso solvente de los recursos de información. Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de especialidad y valorar de forma crítica los resultados de dicha gestión.

CT4. Trabajo en equipo. Ser capaz de trabajar como miembro de un equipo interdisciplinar, ya sea como un miembro más o realizando tareas de dirección, con la finalidad de contribuir a desarrollar proyectos con pragmatismo y sentido de la responsabilidad, asumiendo compromisos teniendo en cuenta los recursos disponibles.

CT7. Tercera lengua. Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

CB4. Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.

CB2. Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio.

### Dedicación: 16h

Actividades dirigidas: 2h

Aprendizaje autónomo: 14h

## SISTEMA DE CALIFICACIÓN

---

Cualquier acto de fraude académico, plagio o uso o mera tenencia al alcance de medios no autorizados en cualquier actividad de evaluación comportará la calificación de cero (0) en la prueba o entrega afectada. Además, de acuerdo con la normativa de la Universidad, la posible derivación de los hechos para la apertura de un expediente disciplinario implicará que la asignatura quede en el estado provisional de "pendiente de evaluación" hasta la resolución del expediente. La gestión de estas incidencias se lleva a cabo de acuerdo con el [Marco de actuación para la integridad académica en la evaluación de la UPC](#).

P = nota de un examen parcial

F = nota de un examen final.

L = nota de los controles de laboratorio (promedio)

La nota final se calculará como  $35\% P + 40\% F + 25\% L$ .

La nota de la competencia transversal "aprendizaje autónomo" se calculará en base a un tema propuesto por el profesor que los alumnos y alumnas tendrán que estudiar por su cuenta. La materia podrá ser objeto de preguntas de examen y/o se requerirá información sobre el tema en alguna de las prácticas de la asignatura.

## BIBLIOGRAFÍA

---

### Básica:

- Baeza-Yates, R.; Ribeiro-Neto, B. Modern information retrieval: the concepts and technology behind search. 2nd ed. Addison-Wesley / Pearson, 2011. ISBN 9780321416919.
- Leskovec, J.; Rajaraman, A.; Ullman, J.D. Mining of massive datasets. 3rd ed. Cambridge: Cambridge University Press, 2020. ISBN 9781108476348.
- Stephens-Davidowitz, S. Everybody lies : what the internet can tell us about who we really are. Bloomsbury Publishing, 2018. ISBN 9781408894736.