

GUÍA DOCENTE ANÁLISIS DE BIG DATA

MÁSTER U. EN INGENIERÍA DE SISTEMAS DE DECISIÓN

CURSO 2024-25

Fecha de publicación: 04-07-2024



I.-Identificación de la Asignatura	
Tipo	OBLIGATORIA
Período de impartición	1 curso, 2S semestre
Nº de créditos	6
Idioma en el que se imparte	Castellano

II.-Presentación
<p>El objetivo de esta asignatura es ofrecer una visión introductoria sobre los principales métodos, tecnologías y herramientas utilizados actualmente para analizar conjuntos de datos complejos y heterogéneos, tales como el denominado <i>big data</i>. Se define <i>big data</i> como aquellos conjuntos de datos cuyo procesamiento y análisis no se puede abordar mediante técnicas de computación o análisis convencionales, precisándose para ello métodos más avanzados. Por ejemplo, con frecuencia implica la utilización de sistemas de computación distribuidos o de muy altas prestaciones.</p> <p>Se describirán los fundamentos teóricos para el tratamiento, procesamiento y análisis de <i>big data</i>, incluyendo aspectos tanto de infraestructura computacional (ingeniería de datos), como de administración de datos (bases de datos no convencionales), así como métodos avanzados de análisis de datos adecuados para este contexto.</p>

III.-Resultados de Aprendizaje

CG01. Adquisición de conocimientos y comprensión de base suficientes para desarrollar y aplicar ideas originales en contextos diversos.

CG02. Aplicación de sus conocimientos y su comprensión, así como sus habilidades para resolver problemas, en entornos nuevos o no familiares y en contextos amplios, empleando para ello información y recursos publicados por profesionales de otros campos y accesibles mediante plataformas de contenidos digitales, archivos, bibliotecas o (si fuese posible) mediante el diálogo directo con dichos profesionales.

CG04. Comunicar y fundamentar sus conclusiones, conocimientos y el marco conceptual en que se basan, de manera clara y sin ambigüedades, tanto a interlocutores expertos como no expertos, utilizando preferentemente para ello tecnologías digitales.

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

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

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

CB9 - Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades

CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

CE12. Capacidad para comprender y aplicar conceptos, métodos y tecnologías fundamentales para el análisis de big data, que permitan incorporarlos a sistemas de decisión avanzados.

CE13. Capacidad para integrar métodos, técnicas y herramientas de diferentes disciplinas (tales como cálculo y procesamiento de datos, estadística, análisis del lenguaje o computación distribuida) para la resolución de problemas específicos.



IV.-Contenido

IV.A.-Temario de la asignatura

Módulo I. Conceptos y métodos para análisis de big data

- Tema 1. Datos estructurados, semi-estructurados y no estructurados.
- Tema 2. Computación paralela y distribuida. MapReduce.
- Tema 3. Tecnologías *cloud computing* y almacenamiento de alto rendimiento.
- Tema 4. Métodos y arquitecturas de procesamiento de datos.

Módulo II. Procesamiento paralelo y distribuido de datos

- Tema 5. Análisis de *big data* en R y Python.
- Tema 6. Ecosistema Apache Hadoop.

Módulo III. Bases de datos no relacionales (NoSQL)

- Tema 7. Metodologías NoSQL.

Módulo IV. Procesamiento de datos en memoria y otras técnicas avanzadas

- Tema 8. Procesamiento de datos en memoria.
- Tema 9. Tecnologías avanzadas de procesamiento distribuido de datos.

Módulo V. Caso práctico

- Tema 10. Resolución de un caso práctico de análisis de datos.

IV.B.-Actividades formativas

Tipo	Descripción
Otras actividades	Tutorías académicas virtuales (plataforma online)
Trabajos colectivos	Desarrollo de trabajos en equipo relacionados con prácticas, seminarios y/o talleres. Participación en eventos programados: foros de debate, tests, chats, etc.
Lecturas	Autoaprendizaje/aprendizaje autónomo, dirigido por el profesor, orientando en el uso del material didáctico necesario a través de la plataforma online.
Lecturas	Preparación de los contenidos teóricos de la asignatura
Resolución de ejercicios	Resolución de preguntas, problemas y ejercicios propuestos a través de la plataforma online

V.-Tiempo de Trabajo del estudiante (30h grado y 25h máster)	
Clases teóricas	0
Clases de resolución de ejercicios, problemas, casos, etc.	44
Prácticas en laboratorios experimentales, tecnológicos, clínicos, campo, etc.	0
Realización de pruebas	4
Tutorías académicas	12
Actividades relacionadas: jornadas, seminarios, etc.	0
Preparación de clases teóricas	40
Preparación de prácticas/ejercicios/casos	40
Preparación de pruebas	10
Total de horas de trabajo del alumnado	150

VI.-Metodología y plan de trabajo		
Tipo	Periodo	Contenido
Tutorías académicas	Semana 1 a Semana 3	Sesiones de resolución de consultas por parte de los alumnos
Lecturas	Semana 1 a Semana 3	Lectura de documentación y referencias propuestas para preparar la resolución de actividades propuestas
Pruebas	Semana 2 a Semana 3	Realización de pruebas de control de nivel de conocimientos asimilados
Prácticas	Semana 1 a Semana 3	Resolución de preguntas, problemas y ejercicios propuestos a través de la plataforma online
Trabajos colectivos	Semana 1 a Semana 3	Trabajos propuestos por el profesor sobre un tema o tecnología específico

VII.-Método de evaluación

El modelo de evaluación general es la evaluación continua, tal como establece el Reglamento de evaluación de los resultados de aprendizaje de la Universidad Rey Juan Carlos.

Deberán utilizarse todos los sistemas de evaluación establecidos para la asignatura en la memoria de la titulación, excepto aquellos que tuviesen una ponderación mínima del 0%, que podrán utilizarse en los cursos académicos en los que el profesorado lo considere oportuno. Cada uno de los sistemas de evaluación podrá ser aplicado mediante una o más actividades de evaluación, coherentes con ese sistema. Ninguna de las actividades de evaluación podrá superar individualmente el 60% de la calificación global de la asignatura.

La suma de las actividades de evaluación no revaluables no podrá superar el 40% de la calificación global de la asignatura y, en general, no deberían tener nota mínima (salvo en el caso de actividades de carácter práctico en las que, estrictamente, no pudieran reproducirse en la convocatoria extraordinaria las condiciones de evaluación de la convocatoria ordinaria).

Los estudiantes que no consigan superar la asignatura en la convocatoria ordinaria, o no se hayan presentado, podrán presentarse a la convocatoria extraordinaria únicamente a las actividades de evaluación revaluables no superadas.

La distribución y características de las actividades de evaluación son las que se describen a continuación.

VII.A.- Descripción de las pruebas de evaluación y su ponderación

Convocatoria ordinaria

La evaluación de la asignatura en convocatoria ordinaria se realizará de acuerdo a las siguientes actividades junto con sus ponderaciones:

- **Resolución de preguntas**, planteadas a través de los foros de la plataforma online. **No hay nota mínima**. Ponderación: 30%. Revaluable en convocatoria extraordinaria.
- **Resolución de casos prácticos y actividades de participación**, a través de la plataforma online. **No hay nota mínima**. Ponderación: 30%. Revaluable en convocatoria extraordinaria.
- **Trabajo individual o en grupo**, presentado a través de la plataforma online. **No hay nota mínima**. Ponderación: 40%. Revaluable en convocatoria extraordinaria.

Para superar **con éxito la asignatura**, es necesario que los alumnos obtengan una **nota global** (suma ponderada de las tres actividades anteriores) **igual o superior a 5.0 sobre 10**.

Convocatoria extraordinaria

Aquellos alumnos que no hayan superado alguna de las tres actividades propuestas deberán presentarse a su reevaluación en convocatoria extraordinaria. Cada una de las tareas tendrá idéntica ponderación a la ya descrita en el apartado de convocatoria ordinaria. En ningún caso se podrá volver a evaluar a un alumno de cualquier tarea que haya superado en convocatoria ordinaria.

Para **superar con éxito la asignatura**, es necesario que los alumnos obtengan una **nota global** (suma ponderada de las tres actividades) **igual o superior a 5.0 sobre 10**.

VII.B.- Evaluación de estudiantes con dispensa académica de asistencia a clase

La concesión de Dispensa Académica de Asistencia a Clase (DAAC no implica que el estudiante quede automáticamente eximido de participar en las actividades de evaluación continua ni en las actividades formativas presenciales de asistencia obligatoria establecidas en la guía docente. Una vez concedida la dispensa, el estudiante deberá contactar con el docente, que podría proponerle las adaptaciones que considere convenientes, siempre que garanticen la adquisición y adecuada evaluación de los resultados de aprendizaje previstos. El estudiante deberá mantener a lo largo de curso una comunicación fluida con el docente para que este le proporcione información sobre las fechas en que se realizarán esas actividades formativas y de evaluación, en caso de que su programación no estuviese ya fijada y a disposición de los estudiantes en el momento de la concesión de la dispensa.

Asignatura con posibilidad de dispensa: Si



VII.C.- Revisión de las pruebas de evaluación

Se realizará conforme al Reglamento de evaluación de los resultados de aprendizaje de la Universidad Rey Juan Carlos.

VII.D.- Estudiantes con discapacidad o necesidades educativas especiales

A fin de garantizar la igualdad de oportunidades, la no discriminación, la accesibilidad universal y la mayor garantía de éxito académico, los y las estudiantes con discapacidad o con necesidades educativas especiales podrán solicitar adaptaciones curriculares para el seguimiento de sus estudios. Esas adaptaciones serán pautadas por la Unidad de Atención a Personas con Discapacidad de la Universidad Rey Juan Carlos, de acuerdo con la normativa que regula el servicio de Atención a Estudiantes con Discapacidad de la Universidad.

Dicha Unidad emitirá un informe de adaptaciones curriculares, por lo que los y las estudiantes con discapacidad o necesidades educativas especiales deberán contactar con la Unidad (discapacidad.programa@urjc.es), a fin de analizar conjuntamente las distintas alternativas.

VII.E.- Conducta académica, integridad y honestidad académica

La Universidad Rey Juan Carlos está plenamente comprometida con los más altos estándares de integridad y honestidad académica, por lo que estudiar en la URJC supone asumir y suscribir los valores de integridad y la honestidad académica recogidos en el Código Ético de la Universidad (<https://www.urjc.es/codigoetico>).

Para acompañar este proceso, la Universidad dispone de la Normativa de Convivencia de la Universidad Rey Juan Carlos (<https://www.urjc.es/images/Universidad/Presentacion/normativa/normativa%20convivencia%20universitaria.pdf>) y de diferentes herramientas (antiplagio, supervisión) que ofrecen una garantía colectiva para el completo desarrollo de estos valores esenciales.



VIII.-Recursos y materiales didácticos	
Bibliografía básica	
Kelleher, J. D., Tierney, B. Data Science . MIT Press, 2018	
Joel Grus. Data Science from Scratch . First principles with Python. 2nd ed. O'Reilly Media, 2019.	
Damji, J.S., Wenig, B., Das, T., Lee, D. Learning Spark: Lightning-Fast Data Analytics . 2nd ed. O'Reilly Media, 2020.	
Provost, F., Fawcett, T. Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking . O'Reilly Media, 2013.	
Bibliografía complementaria	
Harrison, G. Next Generation Databases: NoSQL, NewSQL and Big Data . Apress, 2015.	
Reis, J. y Housley, M. <i>Fundamentals of Data Engineering. Plan and Build Robust Data Systems</i> . O'Reilly Media, 2022.	

IX.-Profesorado	
Nombre y apellidos	JOSE FELIPE ORTEGA SOTO
Correo electrónico	felipe.ortega@urjc.es
Departamento	Teoría de la Señal y Comunicaciones y Sistemas Telemáticos y Computación
Categoría	Profesor/a Contratado/a Doctor/a
Titulación académica	Doctor
Responsable de asignatura	Si
Horario de Tutorías	Para consultar las tutorías póngase en contacto con el/la profesor/a a través de correo electrónico
Nº de Quinquenios	1
Nº de Sexenios	1
Nº de Sexenios de transferencia	0
Nº de evaluaciones positivas Docencia	2
Nombre y apellidos	
ALBERTO FERNANDEZ ISABEL	
Correo electrónico	alberto.fernandez.isabel@urjc.es
Departamento	Informática y Estadística
Categoría	Profesor/a Contratado/a Doctor/a
Titulación académica	Doctor

Responsable de asignatura	No
Horario de Tutorías	Para consultar las tutorías póngase en contacto con el/la profesor/a a través de correo electrónico
Nº de Quinquenios	0
Nº de Sexenios	1
Nº de Sexenios de transferencia	0
Nº de evaluaciones positivas Docencia	1

©2025 Autores Felipe Ortega Soto, Alberto Fernández Isabel

Algunos derechos reservados

Este documento se distribuye bajo la licencia

“Atribución 4.0 Internacional” de Creative Commons, disponible en

<https://creativecommons.org/licenses/by/4.0/deed.es>