

GUÍA DOCENTE

PROCESADO MASIVO DE DATOS

GRADO EN INGENIERÍA BIOMÉDICA (3º y 4º curso)

CURSO 2026-27

Fecha de publicación: 02-07-2026



I.-Identificación de la Asignatura

Tipo	OBLIGATORIA
Período de impartición	4 curso, 2Q semestre
Nº de créditos	3
Idioma en el que se imparte	Castellano

II.-Presentación

La asignatura presenta los conceptos y las herramientas necesarias para utilizar y diseñar sistemas para el análisis de grandes conjuntos de datos, desde la extracción hasta la visualización, aplicados a proyectos de ingeniería biomédica.

De manera más específica, se pretende que el alumno alcance los siguientes objetivos:

1. Entender el contexto histórico de esta disciplina y, por tanto, la importancia y las posibilidades actuales existentes de extraer información, conocimiento y, sobre todo, valor de los datos disponibles tanto en el entorno laboral como especialmente en el ámbito de la biomedicina.
2. Identificar las distintas posibilidades de almacenamiento y de procesamiento de grandes volúmenes de datos, en función de los objetivos establecidos y del tipo de dato disponible.
3. Conocer los distintos tipos de arquitecturas "Big Data" que existen para hacer frente a flujos de datos continuos en función de la velocidad de procesamiento requerida.
4. Aprender las tecnologías disponibles en la actualidad y las cuestiones de diseño necesarias para el almacenamiento y la computación de datos en la nube.
5. Comprender los fundamentos de los modelos de lenguaje masivos, cómo entrenarlos y ajustarlos para fines específicos y cómo evaluarlos, con el fin de manejar datos variados, como el texto libre, especialmente cuando se dispone de grandes cantidades de dichos datos.
6. Asimilar las etapas y procesos del método científico para afrontar de manera rigurosa proyectos de ciencia de datos en el ámbito de la ingeniería biomédica.

Para la correcta consecución de dichos objetivos, se requiere especialmente el conocimiento de la siguiente asignatura:

- Inteligencia Artificial y Aprendizaje

Además, para el máximo aprovechamiento de las clases impartidas y máximo aprendizaje, se requiere también el conocimiento de las siguientes materias:

- Probabilidad y Estadística
- Inferencia
- Optimización
- Fundamentos de Programación
- Bases de Datos
- Análisis de Imágenes Médicas
- Procesado de Señales Fisiológicas



III.-Resultados de Aprendizaje

CG1. Analizar y sintetizar problemas relacionados con la ingeniería biomédica, resolverlos y comunicarlos de forma eficiente.

CG2. Utilizar herramientas computacionales y experimentales para el análisis y cuantificación de problemas de ingeniería biomédica.

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

CE5. Comprender, utilizar y diseñar sistemas de ayuda a la gestión de la información biomédica y a toma de decisiones médicas.

CE6. Conocer, utilizar y diseñar sistemas para el análisis de grandes conjuntos de datos biomédicos, desde la extracción hasta su visualización.

- Leyenda para las titulaciones no adaptadas al RD 822/21: **CB** - competencias básicas, **CG** - competencias generales, **CE** - competencias específicas, **CT** - competencias transversales.
- Leyenda para las titulaciones adaptadas al RD 822/21: **CON** - conocimientos, **COM** - competencias, **HAB** - habilidades.



IV.-Contenido

IV.A.-Temario de la asignatura

Tema 1.— Extraer valor de los datos.

- Principales problemas a resolver.
- De la Inteligencia Empresarial (*Business Intelligence*, BI) a la Ciencia de Datos (*Data Science*, DS).
- Principales perfiles profesionales.

Tema 2.— Almacenar y procesar grandes volúmenes de datos.

- Fuente y tipo de datos. Estructurados, semiestructurados y no estructurados.
- Almacenamiento. Bases de Datos (BBDD) o Almacenes de datos (*Data Warehouse*). Relacional, no relacional o basado en grafos. Veracidad y escalabilidad. SQL, NoSQL o NewSQL.
- Procesamiento. Proceso de Extracción, Transformación y Carga (*Extract, Transform, Load*, ETL). Computación paralela y distribuida. Basada en disco o en memoria: Hadoop o Spark.
- Minería de datos y Aprendizaje máquina (*Machine Learning*, ML) escalable, paralelizable o distribuido.

Tema 3.— Procesamiento de datos a altas velocidades o en tiempo real.

- Procesamiento por lotes (*Batch*) o de flujo (*streaming*).
- Principales arquitecturas *Big Data*.
- *Batch* ML o *streaming* ML.

Tema 4.— Computación en la nube.

- Cuestiones de diseño para el almacenamiento de datos biomédicos y cómputo en la nube.
- Tecnologías actualmente disponibles en la nube.

Tema 5.— Estado actual del *Big Data* y su aplicación en ingeniería biomédica:

- Funcionamiento y evolución de los modelos de lenguaje masivos (*Large Language Models*, LLMs) para procesar datos masivos y heterogéneos. La paralelización de los *transformers* y sus mecanismos de atención como caso de éxito actual.
- Procesamiento *Big Data* mediante técnicas de entrenamiento con datos masivos y heterogéneos. Ajuste fino (*fine-tuning*) de los LLMs para aplicaciones en ingeniería biomédica.
- Evaluación y aplicación de los LLMs. Uso de métricas y *benchmarks*. Desafíos y consideraciones éticas.
- Sistemas basados en LLMs, como Retrieval-Augmented Generation (RAG) y agentes inteligentes, y sus aplicaciones en ingeniería biomédica.

IV.B.-Actividades formativas

Tipo	Descripción
Lecturas	Sesiones teóricas sobre conceptos clave
Prácticas	Resolución de casos prácticos
Laboratorios	Aprendizaje y entrenamiento de herramientas de manejo y análisis de grandes datos
Trabajos colectivos	Reto de la asignatura
Otras actividades	Sesiones de tutorías individuales y por grupos
Laboratorios	Análisis de casos prácticos de Ingeniería Biomédica



V.-Tiempo de trabajo del alumnado (30h grado y 25h máster)

Clases teóricas	10
Clases de resolución de ejercicios, problemas, casos, etc.	18
Prácticas en laboratorios experimentales, tecnológicos, clínicos, campo, etc.	0
Realización de pruebas	2
Tutorías académicas	24
Actividades relacionadas: jornadas, seminarios, etc.	0
Preparación de clases teóricas	15
Preparación de prácticas/ejercicios/casos	15
Preparación de pruebas	6
Total de horas de trabajo del alumnado	90

VI.-Metodología y plan de trabajo

Tipo	Periodo	Contenido
Clases Teóricas	Semana 1 a Semana 15	Clases magistrales.
Trabajos colectivos	Semana 1 a Semana 15	Resolución de casos/proyectos prácticos
Tutorías académicas	Semana 1 a Semana 15	Orientación, seguimiento y retroalimentación individual y grupal.
Pruebas	Semana 6 a Semana 7	Prueba parcial de la parte teórica del curso en formato tipo test.
Prácticas	Semana 1 a Semana 15	Aprendizaje y entrenamiento de herramientas de manejo y análisis de datos.



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).

El alumnado que no consiga 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.

El uso de herramientas basadas en inteligencia artificial está prohibido en cualquier ejercicio y prueba de evaluación de la asignatura salvo que el profesor lo indique de forma explícita en el enunciado o en el material de esta. En caso de que esté permitido, se darán indicaciones precisas de qué usos están permitidos y cuáles no.

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

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

Sistema de Evaluación	Actividad de evaluación	Fecha	Ponderación		Revaluable en Extraordinaria	Nota mínima
SE02 (Examen escrito sobre los conceptos de la asignatura)	Examen escrito (parte teórica)	Semana 6	50%		Sí	3.5
SE05 (Memoria y presentación de prácticas)	Laboratorios/Retos (parte práctica)	De semana 7 a semana 15	Laboratorio/Reto 1: 20%	Total: 50%	Sí	3.5
			Laboratorio/Reto 2: 30%			
Cálculo de la nota final o global • La nota final se calcula como la media ponderada de las notas de las actividades de evaluación según los porcentajes indicados, siempre y cuando se hayan superado con la nota mínima indicada para cada una de ellas. Teniendo en cuenta lo anterior, para aprobar la asignatura la media ponderada debe ser no inferior a 5. • Si una o varias actividades de evaluación no se han superado con la nota mínima, la nota final de la asignatura será la nota media obtenida entre todas las actividades de evaluación realizadas teniendo en cuenta que no podrá superar un 4. • Si no se ha presentado a ninguna prueba evaluable, la nota final será "No presentado" .						
Convocatoria extraordinaria En convocatoria extraordinaria los estudiantes solamente se presentarán a las actividades de evaluación de las pruebas no superadas, de manera que para el cálculo de la nota final en esta convocatoria se utilizará la calificación de las actividades de evaluación aprobadas en convocatoria ordinaria y las notas obtenidas en las actividades revaluadas.						
Convocatoria adelantada El método de evaluación en la convocatoria adelantada será el mismo que el establecido con carácter general para la convocatoria extraordinaria.						
Conducta académica El equipo docente podrá hacer uso de herramientas automáticas de detección de plagio sobre los trabajos presentados. Todos los trabajos entregados de la asignatura deben ser originales, con lo que no podrán utilizarse soluciones realizadas por otras personas distintas a las reflejadas en la autoría del trabajo. Los profesores se reservan la posibilidad de requerir defensa presencial de todos o parte de las/los autoras/es de dichos trabajos si así lo creen necesario. Se recuerda además que, atendiendo al artículo 8.g). de la Normativa de Convivencia de la Universidad Rey Juan Carlos (https://www.urjc.es/images/Universidad/Presentacion/normativa/normativa%20convivencia%20universitaria.pdf) el fraude académico en alguna actividad de evaluación se considera falta muy grave . Las sanciones correspondientes a las faltas muy graves, según el artículo 11 de la referida normativa, son la expulsión temporal de la Universidad, y la pérdida en su caso de los derechos de matrícula.						
VII.B.- Evaluación del alumnado con dispensa académica de asistencia a clase La concesión de <u>Dispensa Académica de Asistencia a Clase (DAAC)</u> no implica que se 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 alumnado 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 alumnado 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 del alumnado en el momento de la concesión de la dispensa. <u>Asignatura con posibilidad de dispensa:</u> Sí						
VII.C.- Revisión y reclamación de las actividades de evaluación Se realizará conforme al Reglamento de evaluación de los resultados de aprendizaje de la Universidad Rey Juan Carlos.						



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

Con el fin de garantizar la igualdad de oportunidades, la no discriminación, la accesibilidad universal y la mayor garantía de éxito académico, las adaptaciones curriculares para alumnado con discapacidad o con necesidades educativas especiales serán pautadas por la Unidad de Atención a Personas con Discapacidad, de acuerdo con la normativa que regula el servicio de Atención a Estudiantes con Discapacidad, aprobada por Consejo de Gobierno de la Universidad Rey Juan Carlos.

Para que esas adaptaciones puedan realizarse, será requisito la emisión de un informe de adaptaciones curriculares por parte de dicha Unidad, por lo que el alumnado con discapacidad o necesidades educativas especiales deberán contactar con este servicio para analizar conjuntamente las distintas opciones.

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.

Para acompañar este proceso, la Universidad dispone de la Normativa de Convivencia de la Universidad Rey Juan Carlos y de diferentes herramientas (antiplagio, supervisión) que ofrecen una garantía colectiva para el completo desarrollo de estos valores esenciales.

Las herramientas basadas en inteligencia artificial pueden mejorar la experiencia de aprendizaje, pero debe utilizarse de forma ética, legal, honesta y transparente. Por este motivo, deben seguirse las recomendaciones de uso adecuado indicadas por los docentes de la asignatura.



VIII.-Recursos y materiales didácticos

Bibliografía básica

Other Readings: Business Culture and Big Data

- * Foster Provost and Tom Fawcett. *Data Science for Business: What you Need to Know about Data Mining and Data-Analytic Thinking*. O'Reilly Media Inc., 2013.
- * Anil Maheshwari. *Big Data Made Accessible: 2019 edition*, and *Data Analytics Made Accessible: 2019 edition*. 2019.
- * John D. Kelleher and Brendan Tierney. *Data Science*. MIT Press Essential Knowledge series, 2018.

Data Visualization

- * Cole Nussbaumer Knaflic. *Storytelling with Data: A Data Visualization Guide for Business Professionals*. John Wiley & Sons, 2015.

Statistical Learning

- * Jerome Friedman, Trevor Hastie, and Robert Tibshirani, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer Series in Statistics, Second Edition, 2009.
(Enlace online: <https://web.stanford.edu/~hastie/Papers/ESLII.pdf>)

Machine Learning

- * Christopher M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- * Andriy Burkov, *The Hundred-Page Machine Learning Book*, 2019.
(Enlace online: <http://themlbook.com/wiki/doku.php>)
- * Tong Xiao y Jingbo Zhu, *Foundations of Large Language Models*. arXiv, 2025.
(Enlace online: <https://arxiv.org/pdf/2501.09223>)

Big Data Systems

- * Martin Kleppmann. *Designing data-intensive applications: The big ideas behind reliable, scalable, and maintainable systems*. O'Reilly Media Inc., 2017.
- * Ralph Kimball and Margy Ross. *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*. John Wiley & Sons, 2002.
- * Nathan Marz and James Warren. *Big Data: Principles and Best Practices of Scalable Real-Time Data Systems*. Manning Publications Co., 2015.

Bibliografía complementaria

IX.-Profesorado

Nombre y apellidos	LUIS BOTE CURIEL
Correo electrónico	luis.bote@urjc.es
Departamento	Teoría de la Señal y Comunicaciones y Sistemas Telemáticos y Computación
Campus de impartición	Fuenlabrada
Categoría	Profesor/a Ayudante 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	0
Nombre y apellidos	SERGIO MUÑOZ ROMERO
Correo electrónico	sergio.munoz@urjc.es
Departamento	Teoría de la Señal y Comunicaciones y Sistemas Telemáticos y Computación
Campus de impartición	Fuenlabrada
Categoría	Titular de Universidad
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	2
Nº de Sexenios	2
Nº de Sexenios de transferencia	0
Nº de evaluaciones positivas Docencia	2