Universidad Rey Juan Carlos

TEACHING GUIDE

MASSIVE DATA PROCESSING

GRADO EN INGENIERÍA BIOMÉDICA (INGLÉS) (3º y 4º curso) (2374)

ACADEMIC YEAR 2026-27

Q2803011B UNIVERSIDAD REY JUAN CARLOS
Fecha firma: 14/11/2026 19:23 | Hash: d2e9bed471836d99de0e4c735844ae1.

Publication date: 02-07-2026

Q2803011B UNIVERSIDAD REY JUAN CARLOS
Fecha firma: 14/11/2026 19:23 | Hash: d2e9bbed471836d99de0e4c735844ae1.

I.-Subject Identification	
Type	OBLIGATORIA
Teaching period	4 course, 2Q semester
Nº of credits	3
Language in wich the subject is taught	English

II.-Presentation

La asignatura presenta los conceptos y las herramientas necesarias para utilizar y disenar sistemas para el analisis de grandes conjuntos de datos, desde la extraccion hasta la visualizacion, 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.-Learning outcomes

CG01 . Analyze and synthesize problems related to biomedical engineering, solve them and communicate them efficiently.
CG02. Use computational and experimental tools for the analysis and quantification of biomedical engineering problems
CB3 . That students have the ability to gather and interpret relevant data (usually within their field of study) to express judgments that include a reflection on relevant topics related to social, scientific or ethical areas.
CE05. Understand, use and design systems to support in the management of biomedical information and medical decision making.
CE06. Know, use and design systems for the analysis of large biomedical data sets, from knowledge extraction to visualization.

Q2803011B UNIVERSIDAD REY JUAN CARLOS
Fecha firma: 14/11/2026 19:23 | Hash: d2e9bed471836d99de0e4c735844ae1.

- Key for degrees not adapted to RD 822/21: **CB** - basic competences, **CG** - generic competences, **CE** - specific competences, **CT** - transversal competences.
- Key for degrees adapted to RD 822/21: **CON** - knowledge, **COM** - competences, **HAB** - skills.

Q2803011B UNIVERSIDAD REY JUAN CARLOS
Fecha firma: 14/11/2026 19:23 | Hash: d2e9bbed471836d99de0e4c735844ae1.

IV.-Contents

IV.A.-Syllabus

Lesson 1.— Extracting value from data.
•Main problems to solve.
•From Business Intelligence (BI) to Data Science (DS).
•Main professional profiles.
Lesson 2.— Storing and processing large data volumes.
•Source and type of data. Structured, semi-structured and unstructured.
•Storage. Databases (BBDD) or Data Warehouses. Relational, non-relational or based on graphs. Veracity and scalability. SQL, NoSQL or NewSQL.
•Processing. Extract, Transform and Load (ETL) process. Parallel and distributed computing. Disk or memory based: Hadoop or Spark.
•Scalable, parallelizable, or distributed data mining (DM) and machine learning (ML).
Lesson 3.— Data processing at high speeds or in real time.
•Batch processing or streaming.
•Main Big Data architectures.
•Batch ML or streaming ML.
Lesson 4.— Cloud computing.
•Design issues for biomedical data storage and cloud computing.
•Currently available cloud technologies.
Topic 5.— Current state of Big Data and its application in Biomedical Engineering
•Operation and evolution of Large Language Models (LLMs) for processing massive and heterogeneous data. Parallelization of transformers and their attention mechanisms as a current success case.
•Big Data processing through training techniques with massive and heterogeneous data. Fine-Tuning of LLMs for applications in Biomedical Engineering.
•Evaluation and application of LLMs. Use of metrics and benchmarks. Challenges and ethical considerations.
•LLM-based systems, such as Retrieval-Augmented Generation (RAG) and intelligent agents, and their applications in biomedical engineering.

IV.B.-Training activities

Type	Title
Theoretical sessions	Theoretical sessions on key concepts.
Resolution of exercises, problems, case studies	Solving practical cases.
Experimental and/or technological laboratories	Learning and Training with management and analysis tools for Biomedical Large Data.
Group work	Course challenge.
Academic tutorials	Individual and group tutorial sessions.

Q2803011B UNIVERSIDAD REY JUAN CARLOS
Fecha firma: 14/11/2026 19:23 | Hash: d2e9bed471836d99de0e4c735844ae1.

V.-Student workload	
Theoretical sessions	10
Resolution of exercises, problems, case studies, etc.	18
Practical sessions in experimental, technological and clinical laboratories, fieldwork, etc.	0
Tests and/or exams	2
Academic tutorials	24
Related activities: conferences, seminars, etc.	0
Preparation of theoretical sessions and self-study	15
Preparation of practical tasks, exercises, case studies, projects, etc.	15
Preparation for tests and/or exams	6
Total student workload	90

VI.-Methodology and academic programme

Type	Period	Content
Tests and/or exams	Week 6 to Week 7	Partial test on the theoretical contents of the course.
Practical sessions	Week 1 to Week 15	Learning and training with tools for Biomedical and Large Data Analysis and other collaborative tools.
Theoretical sessions	Week 1 to Week 15	Masterclasses.
Group work	Week 1 to Week 15	Analyzing practical cases / projects and other collaborative tools.
Academic tutorials	Week 1 to Week 15	Orientation, follow-up, and (individual and group) feedback.

Q2803011B UNIVERSIDAD REY JUAN CARLOS
Fecha firma: 14/11/2026 19:23 | Hash: d2e9bed471836d99de0e4c735844ae1.

Q2803011B UNIVERSIDAD REY JUAN CARLOS
Fecha firma: 14/11/2026 19:23 | Hash: d2e9bed471836d99de0e4c735844ae1.

VII.-Assessment methods

The general assessment method is continuous as established in the *Regulation for the Assessment of Learning Outcomes* of Rey Juan Carlos University.
All the assessment systems established in the degree report must be employed, except for those that have a minimum weight of 0%, which can be used by professors in the academic years they consider relevant. Each assessment system may be applied through one or more assessment activities, consistent with the system. No assessment activity can exceed individually 60% of the overall grade for the subject.
The sum of non re-evaluable activities cannot exceed 40% of the overall grade for the subject and, in general, should not have an established minimum grade (except for practical activities in which, strictly speaking, the conditions of the ordinary call cannot be reproduced in the extraordinary call).
Those students who do not manage to pass the subject in the ordinary call, or do not attend this first call, may attend the extraordinary call only for failed re-evaluable assessment activities.
The distribution and characteristics of the assessment activities are described below.

VII.A.-Description and weight of the assessment activities

Evaluation System	Evaluation Activities	Date	Weight		Can re-eval with extraordinary	Minimum grade
SE02 (Written exam on the subject concepts)	Written exam (theoretical part)	Week 6	50%		Yes	3.5
SE05 (Memo and presentation of practices)	Labs/Challenges (practical part)	Weeks 7 to 15	Lab/Challenge 1: 20% Lab/Challenge 2: 30%	Total: 50%	Yes	3.5

Final grade calculation. The final grade is calculated as the weighted average of the grades obtained in the assessment activities, according to the indicated percentages, provided that the minimum required grade has been achieved for each of them. Based on the above, in order to pass the course, the weighted average must not be lower than 5.
• If one or more assessment activities have not been passed with the minimum required grade, the final grade for the course will be the average of all completed assessment activities, with the condition that it cannot exceed 4.
• If the student has not taken any assessable test, the final grade will be "Not presented".
Extraordinary call. In the extraordinary call, students will only take the assessment activities corresponding to the tests they have not passed. Therefore, for the calculation of the final grade in this session, the grades from the assessment activities passed in the regular session and the grades obtained in the reassessed activities will be used.
Early call. The evaluation method for the early call will be the same as that established in general for the extraordinary call.

VII.B. Assessment of students with academic exemption from class attendance

Having Academic Exemption from Class Attendance (DAAC) does not imply that students are automatically exempt from participating in the continuous assessment activities or in the training activities of compulsory attendance as established in the teaching guide. Once the exemption has been granted, students must contact the subject's teaching staff, who will propose the adaptations he/she may consider relevant, as long as they guarantee the acquisition and adequate assessment of the learning goals established. Students must maintain adequate communication with the teaching staff so that they may provide information on the dates for training and assessment activities, if these are not established in the programme and available for students at the time of achieving the exemption.
Subject with possibility of academic exemption: Yes.

VII.C. Review of assessment activities

In accordance with the Regulation for the Assessment of Learning Outcomes of Rey Juan Carlos University.

VII.D.-Students with a disability or special educational needs

In order to guarantee equal opportunities, non-discrimination, universal accessibility and academic success, students with disabilities or especial education needs may request curricular adaptations to follow their studies. These adaptations will be provided by the Unit of Attention to People with Disabilities of Rey Juan Carlos University, in accordance with the regulation which regulates the service for the Attention of People with Disabilities at the University. This Unit will issue a report for curricular adaptations, so that students with disabilities or special education needs must contact the service (at e-mail discapacidad.programa@urjc.es) in order to provide relevant information for different educational alternatives.

VII.E.-Academic behaviour, academic integrity and honesty

Rey Juan Carlos University is fully committed to the highest standards of academic integrity and honesty, so studying at the URJC means assuming and subscribing to the values of integrity and academic honesty set out in the University's Ethical Code. To support this process, the University has Rey Juan Carlos University's Academic Conduct Regulation as well as different tools (anti-plagiarism, supervision...) that offer a collective guarantee for the full development of these essential values.

Q2803011B UNIVERSIDAD REY JUAN CARLOS
Fecha firma: 14/11/2026 19:23 | Hash: d2e9bed471836d99de0e4c735844ae1.

VIII.-Teaching resources and materials
Basic bibliography
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.
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. (Online link: 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. (Online link: http://themlbook.com/wiki/doku.php) • Tong Xiao y Jingbo Zhu, Foundations of Large Language Models. arXiv, 2025. (Online link: https://arxiv.org/pdf/2501.09223)
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.
Consultation bibliography

IX.-Professors	
Name and surname	LUIS BOTE CURIEL
E-mail address	luis.bote@urjc.es
Department	Teoría de la Señal y Comunicaciones y Sistemas Telemáticos y Computación
Campus	Fuenlabrada
Category	Profesor/a Ayudante Doctor/a
Academic degree	Doctor
Person responsible for the subject	Yes
Tutorial timetable	Para consultar las tutorias póngase en contacto con el/la profesor/a a través de correo electrónico
Number of quinquennia	0
Number of sexennia	1

Q2803011B UNIVERSIDAD REY JUAN CARLOS
Fecha firma: 14/11/2026 19:23 | Hash: d2e9bed471836d99de0e4c735844ae1.

No. of positive evaluations by Docentia	0
Name and surname	SERGIO MUÑOZ ROMERO
E-mail address	sergio.munoz@urjc.es
Department	Teoría de la Señal y Comunicaciones y Sistemas Telemáticos y Computación
Campus	Fuenlabrada
Category	Titular de Universidad
Academic degree	Doctor
Person responsible for the subject	No
Tutorial timetable	Para consultar las tutorias póngase en contacto con el/la profesor/a a través de correo electrónico
Number of quinquennia	2
Number of sexennia	2
No. of positive evaluations by Docentia	2

Q2803011B UNIVERSIDAD REY JUAN CARLOS
Fecha firma: 14/11/2026 19:23 | Hash: d2e9bed471836d99de0e4c735844ae1.