

Escuela de Doctorado y Estudios de Posgrado

Máster Universitario en Ingeniería Informática

GUÍA DOCENTE DE LA ASIGNATURA :

**Análisis de datos masivos
(2025 - 2026)**

1. Datos descriptivos de la asignatura

Asignatura: Análisis de datos masivos	Código: 835941203
<ul style="list-style-type: none">- Centro: Escuela de Doctorado y Estudios de Postgrado- Lugar de impartición: Escuela Superior de Ingeniería y Tecnología. Sección de Ingeniería Informática- Titulación: Máster Universitario en Ingeniería Informática- Plan de Estudios: 2018 (Publicado en)- Rama de conocimiento: Ingeniería y Arquitectura- Ámbito de conocimiento:- Mención/Especialidad:- Itinerario / Intensificación:- Departamento/s: Ingeniería Informática y de Sistemas- Área/s de conocimiento: Ciencia de la Computación e Inteligencia Artificial Lenguajes y Sistemas Informáticos- Curso: 1- Carácter: Obligatoria- Despliegue temporal: Segundo cuatrimestre- Créditos ECTS: 6,0- Modalidad de impartición: Semipresencial- Horario: Enlace al horario- Aula virtual de la asignatura:: http://www.campusvirtual.ull.es- Idioma: Castellano e Inglés (0,30 ECTS en Inglés)	

2. Requisitos de matrícula y calificación

3. Profesorado que imparte la asignatura

Profesor/a Coordinador/a: JOSE MARCOS MORENO VEGA
- Grupo:
General <ul style="list-style-type: none">- Nombre: JOSE MARCOS- Apellido: MORENO VEGA- Departamento: Ingeniería Informática y de Sistemas- Área de conocimiento: Ciencia de la Computación e Inteligencia Artificial

Contacto

- Teléfono 1: **922318175**
- Teléfono 2:
- Correo electrónico: **jmmoreno@ull.es**
- Correo alternativo: **jmmoreno@ull.edu.es**
- Web: **<https://portalciencia.ull.es/investigadores/81201/detalle>**

Tutorías primer cuatrimestre:

Desde	Hasta	Día	Hora inicial	Hora final	Localización	Despacho
Todo el cuatrimestre		Martes	08:00	10:00	Escuela Superior de Ingeniería y Tecnología - Módulo A - AN.4A ESIT	P3.019
Todo el cuatrimestre		Jueves	08:00	12:00	Escuela Superior de Ingeniería y Tecnología - Módulo A - AN.4A ESIT	P3.019

Observaciones:

Tutorías segundo cuatrimestre:

Desde	Hasta	Día	Hora inicial	Hora final	Localización	Despacho
Todo el cuatrimestre		Martes	09:00	11:00	Escuela Superior de Ingeniería y Tecnología - Módulo A - AN.4A ESIT	P3.019
Todo el cuatrimestre		Viernes	08:00	12:00	Escuela Superior de Ingeniería y Tecnología - Módulo A - AN.4A ESIT	P3.019

Observaciones:

Profesor/a: FRANCISCO CARMELO ALMEIDA RODRIGUEZ

- Grupo:

<p>General</p> <ul style="list-style-type: none"> - Nombre: FRANCISCO CARMELO - Apellido: ALMEIDA RODRIGUEZ - Departamento: Ingeniería Informática y de Sistemas - Área de conocimiento: Lenguajes y Sistemas Informáticos 						
<p>Contacto</p> <ul style="list-style-type: none"> - Teléfono 1: 922 31 81 73 - Teléfono 2: - Correo electrónico: falmeida@ull.es - Correo alternativo: - Web: https://portalciencia.ull.es/investigadores/80875/detalle 						
<p>Tutorías primer cuatrimestre:</p>						
Desde	Hasta	Día	Hora inicial	Hora final	Localización	Despacho
Todo el cuatrimestre		Lunes	11:00	13:00	Escuela Superior de Ingeniería y Tecnología - Módulo A - AN.4A ESIT	
Todo el cuatrimestre		Martes	11:00	13:00	Escuela Superior de Ingeniería y Tecnología - Módulo A - AN.4A ESIT	
Todo el cuatrimestre		Miércoles	15:00	17:00	Escuela Superior de Ingeniería y Tecnología - Módulo A - AN.4A ESIT	
<p>Observaciones: Se recomienda solicitar cita previa, especialmente para evitar esperas innecesarias o aglomeraciones. Se podrán desarrollar tutorías en línea previa cita con el profesor. La tutoría se desarrollaría a través de una sesión google meet a desarrollar a través del enlace que se encuentra en el portal de la asignatura para esta actividad.</p>						
<p>Tutorías segundo cuatrimestre:</p>						
Desde	Hasta	Día	Hora inicial	Hora final	Localización	Despacho

Todo el cuatrimestre		Lunes	11:00	13:00	Escuela Superior de Ingeniería y Tecnología - Módulo A - AN.4A ESIT
Todo el cuatrimestre		Martes	11:00	13:00	Escuela Superior de Ingeniería y Tecnología - Módulo A - AN.4A ESIT
Todo el cuatrimestre		Miércoles	15:00	17:00	Escuela Superior de Ingeniería y Tecnología - Módulo A - AN.4A ESIT

Observaciones: Se recomienda solicitar cita previa, especialmente para evitar esperas innecesarias o aglomeraciones. Se podrán desarrollar tutorías en línea previa cita con el profesor. La tutoría se desarrollaría a través de una sesión google meet a desarrollar a través del enlace que se encuentra en el portal de la asignatura para esta actividad.

4. Contextualización de la asignatura en el plan de estudio

Bloque formativo al que pertenece la asignatura: **Tecnologías Informáticas**

Perfil profesional:

5. Resultados de Aprendizaje

Saber (conocimientos o contenidos)

S03 - Conoce el modelado matemático, cálculo y simulación en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación, desarrollo e innovación en todos los ámbitos relacionados con la Ingeniería en Informática

S05 - Conoce y aplica los conocimientos adquiridos para resolver problemas en entornos nuevos o poco conocidos dentro de contextos más amplios y multidisciplinarios, siendo capaces de integrar estos conocimientos

S10 - Comprende, diseña, evalúa y aplica tecnologías, herramientas, recursos, estándares en el marco del desarrollo de software y su implantación en sistemas y entornos libres.

Habilidad (habilidades o destrezas)

H03 - Aprendizaje autónomo, activo y permanente: "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"

H04 - Comunicación oral y escrita: "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"

H05 - Pensamiento crítico y habilidades analíticas: "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"

H06 - Resolución de problemas y toma de decisiones: "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"

H08 - Innovación y creatividad: "Realización, presentación y defensa, una vez obtenidos todos los créditos del plan de estudios, de un ejercicio original realizado individualmente ante un tribunal universitario, consistente en un proyecto integral de Ingeniería en Informática de naturaleza profesional en el que se sinteticen las competencias adquiridas en las enseñanzas"

Competencia

C13 - Capacidad para comprender y poder aplicar conocimientos avanzados de computación de altas prestaciones y métodos numéricos o computacionales a problemas de ingeniería

C15 - Capacidad para aplicar métodos matemáticos, estadísticos y de inteligencia artificial para modelar, diseñar y desarrollar aplicaciones, servicios, sistemas inteligentes y sistemas basados en el conocimiento

6. Contenidos de la asignatura

Contenidos teóricos y prácticos de la asignatura

-Módulo 1: (J. Marcos Moreno Vega)

- Introducción al análisis de datos masivos
- Modelos de clasificación y regresión
- Modelos de agrupamiento
- Reglas de asociación
- Detección de anomalías

Módulo 2: (Francisco Almeida Rodríguez)

- Gestión de datos a gran escala: el paradigma Map-Reduce
- Sistemas de soporte a la gestión de datos a gran escala
- Técnicas y herramientas para la visualización de datos
- Herramientas basadas en software libre para el análisis de datos masivos

Actividades a desarrollar en otro idioma

En virtud de lo dispuesto en la normativa autonómica (decreto 168/2008, 22 de julio) un 5% del contenido será impartido en inglés. En la guía docente de cada asignatura se indicarán las actividades formativas a impartir en inglés y el método de evaluación del contenido correspondiente.

7. Metodología, actividades formativas y volumen de trabajo del estudiante

Modelo de Enseñanza Centrada en el Alumnado

Aplica el Modelo de Enseñanza Centrada en el Alumnado (MECA - ULL)
Aula invertida - Flipped Classroom, Aprendizaje Servicio, Aprendizaje basado en Proyectos (ABP), Aprendizaje basado en Problemas (PBL), Aprendizaje cooperativo, Método o estudio de casos, Simulación

Descripción

* Enseñanza expositiva: clases teóricas donde el profesor expone los conceptos teóricos de la asignatura y clases prácticas o de problemas donde se estudiarán ejemplos prácticos. Se espera que tanto las clases teóricas como las prácticas sean participativas

* Tutorías individuales presenciales o virtuales a través del portal de la asignatura.

* Realización de trabajos/proyectos propuestos por el profesor que serán de tipo práctico y podrán realizarse dentro o fuera del aula virtual, y que serán el material utilizado para la evaluación continua.

* Dentro de las actividades on-line que se realizan en la asignatura, vamos a distinguir dos modalidades:

a) Online asíncrona en las que las comunicaciones profesor/alumno se realizan mediante herramientas de mensajería (correo electrónico, foros de discusión, ...). En estas actividades, se propondrá un aspecto teórico/práctico de la asignatura que el alumno deberá abordar haciendo uso del material suministrado por el profesor. Mediante sistemas de mensajería se establecerá el canal para la resolución de dudas o discusiones y se propondrá el entregable asociado (máquina virtual configurada, cuestionario,...)

b) Online síncrona en las que se fija un horario en el que se establece una comunicación directa profesor/alumno usando herramientas como chats y videoconferencias

Actividades formativas y horas de trabajo del estudiante

Actividades formativas	Horas	Presencialidad [%]
Clases teóricas	6,00	100,00
Clases prácticas en aula de informática o laboratorio	30,00	50,00
Seminarios-Talleres u otras actividades complementarias (foros de debate, simulaciones, ...)	12,00	25,00
Trabajo autónomo	90,00	0,00
Exposiciones, debates, y defensas de trabajos y proyectos	9,00	25,00
Evaluación.	3,00	100,00

Total horas y presencialidad	150,00	48,75
Total ECTS	6,00	

8. Bibliografía / Recursos

Bibliografía Básica

José Hernández Orallo, M.José Ramírez Quintana, Cesar Ferri Ramírez.

Introducción a la Minería de Datos

. Editorial Pearson, 2004

Ethem Alpaydin,

Introduction to Machine Learning

, The MIT Press, 2004

Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, Vipin Kumar,

Introduction to Data Mining

, Addison-Wesley, 2006

Charu. C. Aggarwal,

Data Mining.

The Textbook

. Springer, 2015

Bibliografía Complementaria

Ian H. Witten, Eibe Frank, Mark A. Hall, Christopher J. Pal,

Data Mining

.

Practical Machine Learning Tools and Techniques

, Morgan Kaufmann, 2017

Foster Provost, Tom Fawcett, Data Science for Business, O'Reilly Media, 2013

Michael R. Berthold, Christian Borgelt, Frank Höppner, Frank Klawonn,

Guide to Intelligent Data Analysis

, Springer, 2010

Yanchang Zhao,

R and Data Mining: Examples and Case Studies,

Academic Press, 2012

Michael J. Crawley, The R Book, Wiley, 2013

Otros Recursos

R project

kaggle

Google colab

Python

Scipy

Apache Spark

9. Sistema de evaluación y calificación

Descripción

La evaluación de la asignatura se rige por el Reglamento de Evaluación y Calificación de la Universidad de La Laguna, aprobado el 21 de junio de 2022, modificado por acuerdos del Consejo de Gobierno de 13 de julio de 2022, 8 de noviembre de 2022 y 31 de mayo de 2023, además de por lo establecido en la Memoria de Verificación inicial del título o posteriores modificaciones.

Atendiendo a lo establecido en los Estatutos de la Universidad de La Laguna, la modalidad de evaluación continua será obligatoria en la primera convocatoria de la asignatura para todo el alumnado (excepto para quienes se acojan a la evaluación única según se dispone en el artículo 5.5 del Reglamento de Evaluación y Calificación de la Universidad de La Laguna).

Podrán acogerse a la modalidad de evaluación única, en la primera convocatoria, quienes lo comuniquen, a través del procedimiento habilitado en el aula virtual de la asignatura, antes de haberse presentado a las actividades cuya ponderación compute el 40% de la evaluación continua.

Una convocatoria se entenderá agotada cuando el/la alumno/a se presente, al menos, a las actividades cuya ponderación compute el 50% de la evaluación continua.

El alumnado que no haya superado la asignatura en la primera convocatoria, o el que se haya acogido a la modalidad de evaluación única, deberá examinarse de todas las actividades obligatorias de la evaluación continua que no haya superado en dicha convocatoria.

El alumnado que se encuentre en la quinta o posteriores convocatorias y desee ser evaluado por un Tribunal, deberá presentar una solicitud a través del procedimiento habilitado en la sede electrónica, dirigida al Decanato de la Escuela. Dicha solicitud deberá realizarse con una antelación mínima de diez días hábiles al comienzo del periodo de exámenes.

Es obligatorio asistir a clases y hacer uso de los foros y tutorías tanto presenciales como on-line.

Evaluación continua.-

La evaluación continua tiene en cuenta las siguientes actividades:

- Prácticas individuales a desarrollar en los laboratorios frente al ordenador, pruebas teóricas y/o online a través del campus virtual. Se contemplan pruebas para los módulos de análisis de datos (módulo 1) y visualización (módulo 2).
- Proyectos prácticos de análisis (módulo 1) y visualización de datos (módulo 2).

Todas las actividades de la evaluación continua tienen carácter obligatorio.

Evaluación única.-

Consistirá en la entrega y defensa oral de los proyectos y en la realización o entrega de las prácticas individuales anteriores en las fechas establecidas para ello por la Universidad de La Laguna y la Escuela Superior de Ingeniería y Tecnología.

Si el alumnado así lo manifiesta, se trasladará la calificación de las actividades obligatorias que haya superado en la evaluación continua a la evaluación única. En ningún caso, existe la posibilidad de presentarse a una actividad obligatoria ya aprobada para subir nota. Tampoco se mantendrá la calificación obtenida en las diferentes pruebas evaluativas de un curso al siguiente.

El software utilizado en las prácticas de laboratorio, así como los manuales y tutoriales que usarán los alumnos para el desarrollo de las mismas, están documentados en inglés. Se evaluará la comprensión de los mismos en la calificación del proyecto y de los ejercicios teórico prácticos.

La ponderación de las actividades en la calificación final será de 50% para el proyecto y pruebas del módulo 1 y 50% para el proyecto y pruebas del módulo 2. Debe superarse las actividades asociadas a cada módulo por separado para superar la asignatura.

Estrategia Evaluativa

Pruebas evaluativas	Competencias	Criterios	Ponderación
Evaluación de Teoría	[S10], [H06], [H05], [C15], [H08], [H04], [H03], [S05], [S03], [C13]	<ul style="list-style-type: none"> • Estructura, calidad y claridad de la redacción, rigor en el análisis de los datos y coherencia de las respuestas dadas a las pruebas evaluativas de los contenidos teóricos. • Grado de conocimiento adquirido en las herramientas y técnicas para el análisis y visualización de los datos. 	40,00 %

Evaluación de prácticas, trabajo y proyectos	[S10], [H06], [H05], [C15], [H08], [H04], [H03], [S05], [S03], [C13]	<ul style="list-style-type: none"> • Pertinencia de las herramientas usadas para el análisis y la visualización de los datos. • Adecuación de las respuestas dadas a las prácticas y claridad de redacción de los informes de las mismas. 	60,00 %
--	--	---	---------

10. Cronograma / calendario de la asignatura

Descripción

-El cronograma que se presenta a continuación describe las actividades que se van a desarrollar durante el cuatrimestre en el que se imparte la asignatura. Sin embargo, esta planificación puede ser modificada si así lo demanda el desarrollo de la asignatura.

En dicho cronograma se presentan actividades que son presenciales (A) y otras que se desarrollan de forma online utilizando las herramientas TIC disponibles. Dentro de estas actividades "online", se indicará con "C" aquellas actividades online asíncronas y con "B", aquellas que sean síncronas.

Debido al carácter semipresencial del máster habrá semanas en las que las clases prácticas, trabajos y proyectos se desarrollarán en formato no-presencial según lo dispuesto en el horario de clases del máster.

Segundo cuatrimestre

Semana	Temas	Actividades de enseñanza aprendizaje	Horas de trabajo presencial	Horas de trabajo autónomo	Total
Semana 1:	Introducción al análisis de datos masivos	Clase teórica. Laboratorio online (C)	2.00	8.00	10.00
Semana 2:	Modelos de clasificación y regresión	Clase teórica. Laboratorio presencial (A)	2.00	7.00	9.00
Semana 3:	Modelos de clasificación y regresión	Clase teórica. Laboratorio online (B)	2.00	7.00	9.00
Semana 4:	Modelos de agrupamiento	Clase teórica. Laboratorio online (C)	2.00	8.00	10.00

Semana 5:	Modelos de agrupamiento	Clase teórica. Laboratorio presencial (A)	2.00	8.00	10.00
Semana 6:	Reglas de asociación	Clase teórica. Laboratorio online (B)	2.00	7.00	9.00
Semana 7:	Detección de anomalías	Clase teórica. Laboratorio online (C)	2.00	7.00	9.00
Semana 8:	Gestión de datos a gran escala: el paradigma Map-Reduce	Clase teórica. Laboratorio presencial (A)	2.00	8.00	10.00
Semana 9:	Sistemas de soporte a la gestión de datos a gran escala	Clase teórica. Laboratorio online (B)	2.00	7.00	9.00
Semana 10:	Sistemas de soporte a la gestión de datos a gran escala	Clase teórica. Laboratorio online (C)	2.00	8.00	10.00
Semana 11:	Técnicas y herramientas para la visualización de datos	Clase teórica. Laboratorio presencial (A)	2.00	7.00	9.00
Semana 12:	Técnicas y herramientas para la visualización de datos	Clase teórica. Laboratorio online (B)	2.00	8.00	10.00
Semana 13:	Técnicas y herramientas para la visualización de datos	Clase teórica. Laboratorio online (C)	2.00	7.00	9.00
Semana 14:	Herramientas basadas en software libre para el análisis de datos masivos	Clase teórica. Laboratorio presencial (A)	2.00	8.00	10.00
Semana 15:	Presentación de los proyectos		1.00	8.00	9.00
Semana 16 a 18:	Presentación de los proyectos		1.00	7.00	8.00
Total			30.00	120.00	150.00