

GUÍA DOCENTE

DENOMINACIÓN DE LA ASIGNATURA

Denominación: **INTRODUCCIÓN AL BIG DATA ANALYTICS(BDA)**
Código: 634006
Plan de estudios: **MÁSTER UNIVERSITARIO EN INTELIGENCIA COMPUTACIONAL E INTERNET DE LAS COSAS** Curso: 1
Créditos ECTS: 4.0 Horas de trabajo presencial: 30
Porcentaje de presencialidad: 30.0% Horas de trabajo no presencial: 70
Plataforma virtual: <http://www.uco.es/moodle>

DATOS DEL PROFESORADO

Nombre: GARCIA MARTÍNEZ, CARLOS (Coordinador)
Departamento: INFORMÁTICA Y ANÁLISIS NUMÉRICO
Área: CIENCIA DE LA COMPUTACIÓN E INTELIGENCIA ARTIFICIAL
Ubicación del despacho: Dpto. Informática y Análisis Numérico. Anexo C3 en Edificio Marie Curie, planta baja (X3BN090)
E-Mail: cgarcia@uco.es Teléfono: 957 21 26 60

Nombre: LUNA ARIZA, JOSÉ MARÍA
Departamento: INFORMÁTICA Y ANÁLISIS NUMÉRICO
Área: CIENCIA DE LA COMPUTACIÓN E INTELIGENCIA ARTIFICIAL
Ubicación del despacho: Dpto. Informática y Análisis Numérico. Edificio Albert Einstein, planta tercera
E-Mail: jmluna@uco.es Teléfono: +34957212218
URL web: <http://www.uco.es/grupos/kdis/www/members/jmluna.html>

Nombre: RAMÍREZ QUESADA, AURORA
Departamento: INFORMÁTICA Y ANÁLISIS NUMÉRICO
Área: CIENCIA DE LA COMPUTACIÓN E INTELIGENCIA ARTIFICIAL
Ubicación del despacho: Dpto. Informática y Análisis Numérico. Anexo C3 en Edificio Marie Curie, planta baja (X3BN080)
E-Mail: aramirez@uco.es Teléfono: +34957212660
URL web: <https://www.uco.es/users/aramirez/>

Nombre: ROMERO SALGUERO, JOSÉ RAÚL
Departamento: INFORMÁTICA Y ANÁLISIS NUMÉRICO
Área: CIENCIA DE LA COMPUTACIÓN E INTELIGENCIA ARTIFICIAL
Ubicación del despacho: Dpto. Informática y Análisis Numérico. Anexo C3 en Edificio Marie Curie, planta baja (X3BN100)
E-Mail: jrromero@uco.es Teléfono: +34957212660

REQUISITOS Y RECOMENDACIONES

Requisitos previos establecidos en el plan de estudios

Ninguno

Recomendaciones

Conocimientos de bases de datos relacionales y lenguajes de programación (Python o Java).

GUÍA DOCENTE

COMPETENCIAS

CG1	Obtener información, diseñar experimentos e interpretar resultados en los ámbitos de la Inteligencia Computacional y el Internet de las Cosas
CG2	Manejar las fuentes de información científica y recursos útiles para el estudio y la investigación en los ámbitos de la Inteligencia Computacional y el Internet de las cosas
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 ⁹ relacionados con su área de conocimiento.
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.
CT1	Analizar y sintetizar conocimiento y realizar un razonamiento crítico.
CT3	Aplicar los conocimientos adquiridos en la resolución de problemas en contextos nuevos.
CE2	Analizar y modelar grandes volúmenes de datos, estructurados y no estructurados, generando propuestas estratégicas e innovadoras mediante la aplicación de técnicas de análisis de datos.
CE5	Modelar y resolver problemas reales o académicos mediante técnicas de ciencia de datos.

OBJETIVOS

- + Conocer los conceptos básicos e identificar las tecnologías asociadas al campo de Big Data
- + Conocer la problemática asociada al manejo y salvaguarda de grandes volúmenes de datos, así como las características inherentes a este tipo de problemas.
- + Analizar qué tecnología se adapta mejor a las necesidades de un problema real de volúmenes de datos masivos.
- + Conocer los métodos y tecnologías para Big Data Analytics, con especial énfasis en técnicas de aprendizaje automático.

CONTENIDOS

1. Contenidos teóricos

- + Introducción
- + Almacenamiento en sistemas de *big data*
- + Procesamiento en sistemas de *big data*
- + Ecosistema Hadoop
- + Analítica de datos en sistemas de *big data*.

2. Contenidos prácticos

- + Instalación y administración de Hadoop y el sistema de archivos HDFS.
- + Introducción al paradigma de computación Map-Reduce
- + Herramientas del ecosistema Hadoop.
- + Introducción al sistema Spark y a MLlib

OBJETIVOS DE DESARROLLO SOSTENIBLE RELACIONADOS CON LOS CONTENIDOS

Sin relación

GUÍA DOCENTE**METODOLOGÍA****Aclaraciones**

El alumnado a tiempo parcial deberá estar al tanto de los registros y avances publicados en la página de la asignatura.

Actividades presenciales

Actividad	Total
<i>Actividades de evaluación</i>	2.5
<i>Laboratorio</i>	7.5
<i>Lección magistral</i>	15
<i>Trabajos en grupo (cooperativo)</i>	2.5
<i>Tutorías</i>	2.5
Total horas:	30.0

Actividades no presenciales

Actividad	Total
<i>Búsqueda de información</i>	20
<i>Ejercicios</i>	20
<i>Estudio</i>	30
Total horas:	70

MATERIAL DE TRABAJO PARA EL ALUMNO

Ejercicios y problemas
Notebooks de código Python
Presentaciones PowerPoint

Aclaraciones

Toda la información necesaria será puesta a disposición del alumnado en la plataforma MOODLE.

EVALUACIÓN

Instrumentos	Porcentaje
Cuestionarios on-line	25%
Informes/memorias de prácticas	60%
Registros de observación	15%

GUÍA DOCENTE

Periodo de validez de las calificaciones parciales:

Convocatorias del mismo curso académico.

Aclaraciones:

Para la evaluación de la parte teórica, se utilizarán una serie de cuestionarios presentados durante el transcurso de la asignatura.

Para la evaluación de la parte práctica, se realizarán entregas de las actividades realizadas durante el transcurso de la asignatura. Se calculará la media ponderada de las calificaciones asignadas a dichas prácticas, teniendo mayor peso aquellas de mayor complejidad.

Se valorará adicionalmente la entrega, de memorias y realización de cuestionarios, en los plazos inicialmente fijados, según indica la tabla en el instrumento Registros de observación.

Las convocatorias extraordinarias y extraordinaria de fin de estudios se registrará por las mismas indicaciones que las convocatorias ordinarias.

Aclaraciones:

BIBLIOGRAFIA

1. Bibliografía básica

+ B. Balamurugan, et al. Big Data: Concepts, Technology, and Architecture. Wiley, 2021. ISBN: 978-1-119-70182-8

+ A. Maheshwari. Big Data (2nd edition). Mc Graw Hill, 2019. ISBN: 9353167957

+ Big Data Made Easy - A Working Guide to the Complete Hadoop Toolset. Apress 2015. ISBN: 978-1484200957

+Tom White. Hadoop The Definitive Guide. 4th ed. O'Reilly. 2015.

Además de las referencias bibliográficas indicadas, se proporcionará al estudiante referencias y lectura actualizada del tema a través de la plataforma Moodle.

CC2020: Computing Curricula 2020: Paradigms for Global Computing Education.

2. Bibliografía complementaria

+ V. Naresh Kumar, P. Shindgikar. Modern Big Data Processing with Hadoop. Packt, 2018. ISBN. 1787122276X

+ S. Ryza, U. Laserson, S. Owen, J. Wills. Advanced Analytics with Spark. O'Reilly, 2017. ISBN: 1491972953

+ W. McKinney. Python for Data Analysis, 2e: Data Wrangling with Pandas, Numpy and IPython. O'Reilly, 2017. ISBN: 1491957662

+ R. Mehta. Big Data Analytics with Java. Packt, 2017. ISBN: 1787288980

Las estrategias metodológicas y el sistema de evaluación contempladas en esta Guía Docente serán adaptadas de acuerdo a las necesidades presentadas por estudiantes con discapacidad y necesidades educativas especiales en los casos que se requieran.