



COMILLAS

UNIVERSIDAD PONTIFICIA

ICAI

ICADE

CIHS

GUÍA DOCENTE

2024 - 2025

FICHA TÉCNICA DE LA ASIGNATURA

Datos de la asignatura	
Nombre completo	Tecnologías de procesamiento Big Data
Código	DTC-MBD-515
Título	Máster en Big Data. Tecnología y Analítica Avanzada/Master in Big Data Technologies and Advanced Analytics
Impartido en	Máster en Big Data. Tec. y Analítica Avanzada/Master in Big Data Technologies and Advanced Analytics [Primer Curso]
Nivel	Master
Cuatrimestre	Semestral
Créditos	7,5 ECTS
Carácter	Obligatoria
Departamento / Área	Departamento de Telemática y Computación

Datos del profesorado	
Profesor	
Nombre	Patricia Alcalde Sanz
Departamento / Área	Departamento de Telemática y Computación
Correo electrónico	palcalde@icai.comillas.edu
Profesor	
Nombre	Jorge Ayuso Rejas
Departamento / Área	Departamento de Telemática y Computación
Correo electrónico	jayuso@icai.comillas.edu
Profesor	
Nombre	José Ramón Fernández Acosta
Departamento / Área	Departamento de Telemática y Computación
Correo electrónico	jrfacosta@icai.comillas.edu

DATOS ESPECÍFICOS DE LA ASIGNATURA

Contextualización de la asignatura
Aportación al perfil profesional de la titulación
<p>El objetivo del curso es dar una visión general de las herramientas de ingesta y procesamiento del entorno big data, especialmente centrado en Spark y Hive.</p> <p>Al finalizar el curso, los alumnos:</p>



- Ser capaces de elegir cuál es la herramienta más adecuada para extraer datos de diferentes fuentes y llevarlos a un cluster Hadoop.
- Tener experiencia con algunas herramientas y lenguajes de procesamiento (python, sql, etc).
- Tener conocimientos profundos de spark con python y de cómo optimizar trabajos.

Prerrequisitos

Los estudiantes que deseen realizar este curso deben estar familiarizados con cualquier lenguaje de programación, preferiblemente python o SQL, y con los comandos y utilidades de Linux.

Competencias - Objetivos

Competencias

Competencias

CP1 Integrar las arquitecturas, técnicas de inteligencia artificial, análisis avanzado de datos y de visualización y de cumplimiento legal para ofrecer la solución global óptima.

CP2 Aplicar e integrar los flujos programáticos de datos masivos.

CP4 Implementar las técnicas de procesado de datos y usar las herramientas más habituales y apropiadas a las condiciones y requisitos de casos específicos

CP7 Aplicar conocimientos avanzados en Big Data y analítica de datos para desarrollar soluciones innovadoras en proyectos y en investigación, aportando y evaluando soluciones óptimas para el procesamiento y análisis de datos a gran escala

Habilidades o destrezas

HA1 Comunicar de manera oral y escrita con rigor técnico, claridad expositiva y coherencia argumentativa a todo tipo de interlocutores, técnicos y no técnicos.

HA2 Trabajar en equipos de carácter pluridisciplinar y/o internacional y organizar y liderar adecuadamente las dinámicas de grupo.

HA3 Desarrollar las habilidades interpersonales que requieren los entornos profesionales actuales (empatía, tolerancia, respeto, capacidad para aunar intereses contrapuestos).

HA4 Gestionar, organizar y planificar adecuadamente el trabajo y el tiempo, cumpliendo objetivos y estándares de calidad.

HA5 Mantener una formación y aprendizaje continuo y adaptación a los cambios tecnológicos y científicos.

Resultados de Aprendizaje

Conocimientos o contenidos

CO1 Entender los fundamentos de la analítica de datos y su aplicación en diversas áreas de la inteligencia artificial, destacando la integración en soluciones complejas y multidisciplinares para el análisis avanzado de datos masivos atendiendo a la diversidad de problemas específicos de cada área.

CO2 Comprender las técnicas de procesados de datos, las arquitecturas y herramientas más habituales y apropiadas para condiciones y requisitos de casos específicos.

BLOQUES TEMÁTICOS Y CONTENIDOS

Contenidos – Bloques Temáticos

1. Configuración del entorno. IntelliJ y Git.
2. Sistemas de almacenamiento distribuidos de datos estructurados y no estructurados.



3. Pipelines de ingesta de grandes cantidades de datos en entornos distribuidos.
4. Sistemas de mensajería mediante colas de eventos.
5. Python Científico, principales librerías para utilizar Python en el ámbito científico y en el procesamiento de datos.
6. Motores de procesamiento Big Data. Arquitectura. Diseño de aplicaciones. Rendimiento.
7. Introducción a los motores de búsqueda.
8. Herramientas de ETL.

METODOLOGÍA DOCENTE

Aspectos metodológicos generales de la asignatura

Metodología Presencial: Actividades

Clases teóricas enfocadas a adquirir los conocimientos necesarios para realizar las prácticas.

Ejercicios resueltos en clase guiados por el profesor.

Ejercicios grupales de resolución en clase para que trabajen en equipo .

Metodología No presencial: Actividades

Propuesta de prácticas de realización individual de cada uno de los módulos. Estas prácticas formarán parte de la nota final de la asignatura.

RESUMEN HORAS DE TRABAJO DEL ALUMNO

Actividad Formativa	Horas	%
Clases magistrales expositivas y participativas	35	100
Ejercicios prácticos y resolución de problemas	40	100
Estudio personal	40	0
Proyectos	60	0
Trabajos	50	0

EVALUACIÓN Y CRITERIOS DE CALIFICACIÓN

Sistema de evaluación	Ponderación mínima	Ponderación máxima
Examen escrito/oral/test	40%	40%
Trabajo/Proyecto/Caso práctico	60%	60%



COMILLAS

UNIVERSIDAD PONTIFICIA

ICAI

ICADE

CIHS

GUÍA DOCENTE

2024 - 2025

individual/grupo

BIBLIOGRAFÍA Y RECURSOS

Bibliografía Básica

- Apuntes y notebooks preparados por el profesor (disponibles en Moodle).
- White, T. (2015). Hadoop: The definitive guide 4th edition. " O'Reilly Media, Inc."
- Shreedharan, Hari (2014). Using Flume " O'Reilly Media, Inc."
- Karau, H., Konwinski, A., Wendell, P., & Zaharia, M. (2015). Learning spark: lightning-fast big data analysis. " O'Reilly Media, Inc."
- VanderPlas, J. (2016). Python Data Science Handbook.

En cumplimiento de la normativa vigente en materia de **protección de datos de carácter personal**, le informamos y recordamos que puede consultar los aspectos relativos a privacidad y protección de datos que ha aceptado en su matrícula entrando en esta web y pulsando "descargar"

<https://servicios.upcomillas.es/sedelectronica/inicio.aspx?csv=02E4557CAA66F4A81663AD10CED66792>