



## FICHA TÉCNICA DE LA ASIGNATURA

Datos de la asignatura	
Nombre completo	Análisis de datos no estructurados
Código	DTC-MBD-522
Título	Máster en Big Data. Tecnología y Analítica Avanzada/Master in Big Data Technologies and Advanced Analytics
Impartido en	Máster Universitario en Ingeniería de Telecomunicación + Máster Big Data.Tecnología y Anal. Avanzada [Segundo Curso] Máster en Big Data. Tec. y Analítica Avanzada/Master in Big Data Technologies and Advanced Analytics [Primer Curso]
Nivel	Master
Cuatrimestre	Semestral
Créditos	6,0 ECTS
Carácter	Optativa
Departamento / Área	Departamento de Telemática y Computación
Responsable	Ana Laguna Pradas

Datos del profesorado	
<b>Profesor</b>	
Nombre	Ana Laguna Pradas
Departamento / Área	Departamento de Telemática y Computación
Correo electrónico	alaguna@icai.comillas.edu
<b>Profesores de laboratorio</b>	
<b>Profesor</b>	
Nombre	Cristina Puente Águeda
Departamento / Área	Departamento de Telemática y Computación
Despacho	D-408 ICAI Alberto Aguilera 25
Correo electrónico	cristina.puente@icai.comillas.edu

## DATOS ESPECÍFICOS DE LA ASIGNATURA

Contextualización de la asignatura
<b>Aportación al perfil profesional de la titulación</b>
<p>En el contexto del Master de Big Data: Tecnología y analítica avanzada, esta asignatura pretende aportar una visión avanzada al tratamiento de datos no estructurados como son las imágenes, el sonido, logs o documentos de texto.</p> <p>El alumno será capaz de tener una visión completa y general las técnicas más avanzadas de tratamiento de datos.</p>



## Prerrequisitos

Se requiere fundamentos de programación, lenguaje python y tratamiento de datos.

## 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.

CP3 Implementar técnicas de análisis estadístico y lenguajes de programación para Machine Learning en el contexto de Big Data, ajustando la metodología a las especificidades de cada conjunto de datos para optimizar los resultados obtenidos.

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

CP5 Aplicar los principios éticos relativos a la recogida, almacenamiento, y análisis de datos teniendo en cuenta las posibles discriminaciones directas o indirectas derivadas de la toma de decisiones.

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. Breve descripción de los contenidos de la materia

## Resultados de Aprendizaje

### Resultados del proceso de formación y 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



CO3 Comprender las técnicas de machine Learning e Inteligencia artificial, desde las más sencillas a las más sofisticadas y ser capaz de seleccionar la más adecuada e implementarla de una manera adecuada a las condiciones y requisitos del caso específico.

## BLOQUES TEMÁTICOS Y CONTENIDOS

### Contenidos – Bloques Temáticos

#### Contenidos – Bloques Temáticos

##### Section 1: LOGS

Intro to Data Logs Analysis

Acquiring Data Logs

Pre-processing Data Logs

##### Section 2: IMAGE

Intro to Computer Vision

Basic Image Processing

Handcrafted feature extraction and Machine Learning

Deep Learning (CNNs, Data augmentation, Transfer learning)

Transfer learning (Feature extraction and Fine tuning)

Advanced Deep Learning: Feature maps Visualization

Generative Deep Learning, Stable diffusion models

##### Section 3: AUDIO

Intro to Sound Data

Audio Data Pre-processing (Sound properties, Audio data characteristics, etc.)

Feature Extraction (MFCCs, Spectrograms, etc.)

Data Exploration and Audio Visualization (Clustering, etc.)

Deep Learning with Audio data

Generative Deep Learning, Speech-to-Text, Text-to-Audio

##### Section 4: TEXT



Intro to Natural Language Processing (tokenization, lemmatization, POS, NER, etc.)

Feature extraction. One hot encoding. Document representation

Dimensionality Reduction. Words Embeddings. Machine Learning with NLP

Deep Learning applied to sequence data (RNN, LSTM, etc.)

Generative Deep Learning, Transformers, GPT, Reinforcement Learning

Section 5: GRAPHS

Intro to Network Analysis and Graph theory

## METODOLOGÍA DOCENTE

### Aspectos metodológicos generales de la asignatura

#### Metodología Presencial: Actividades

Metodología Presencial: Actividades

1. Clase magistral y presentaciones generales. Exposición de los principales conceptos y procedimientos mediante la explicación por parte del profesor. Incluirá presentaciones dinámicas, pequeños ejemplos prácticos y la participación reglada o espontánea de los estudiantes (25 horas).
2. Prácticas de laboratorio. Las prácticas de laboratorio podrán requerir la realización de un trabajo previo de preparación y finalizar con la redacción de un informe de laboratorio o la inclusión de las distintas experiencias en un cuaderno de laboratorio (35 horas).
3. Tutorías. Se realizarán en grupo e individualmente para resolver las dudas que se les planteen a los alumnos después de haber trabajado los distintos temas. Y también para orientar al alumno en su proceso de aprendizaje

#### Metodología No presencial: Actividades

Metodología No presencial: Actividades

El objetivo principal del trabajo no presencial es llegar a entender y comprender los conceptos teóricos de la asignatura, así como ser capaz de poner en práctica estos conocimientos para resolver los diferentes tipos de problemas

1. Estudio individual del material a discutir en clases posteriores. Actividad realizada individualmente por el estudiante cuando analiza, busca e interioriza la información que aporta la materia y que será discutida con sus compañeros y el profesor en clases posteriores.
2. Prácticas de laboratorio. Las prácticas de laboratorio podrán requerir la realización de un trabajo previo de preparación y finalizar con la redacción de un informe de laboratorio o la inclusión de las distintas experiencias en un cuaderno de laboratorio.

## RESUMEN HORAS DE TRABAJO DEL ALUMNO

#### Metodología Presencial: Actividades

Clase magistral y presentaciones generales. Exposición de los principales conceptos y procedimientos mediante la explicación por parte del profesor. Incluirá presentaciones dinámicas, pequeños ejemplos prácticos y la participación reglada o espontánea de los estudiantes



(25 horas).

Prácticas de laboratorio. Las prácticas de laboratorio podrán requerir la realización de un trabajo previo de preparación y finalizar redacción de un informe de laboratorio o la inclusión de las distintas experiencias en un cuaderno de laboratorio (35 horas).

Tutorías. Se realizarán en grupo e individualmente para resolver las dudas que se les planteen a los alumnos después de haber trabajado los distintos temas. Y también para orientar al alumno en su proceso de aprendizaje

### Metodología No presencial: Actividades

El objetivo principal del trabajo no presencial es llegar a entender y comprender los conceptos teóricos de la asignatura, así como capaz de poner en práctica estos conocimientos para resolver los diferentes tipos de problemas

Estudio individual del material a discutir en clases posteriores. Actividad realizada individualmente por el estudiante cuando busca e interioriza la información que aporta la materia y que será discutida con sus compañeros y el profesor en clases posteriores.

Prácticas de laboratorio. Las prácticas de laboratorio podrán requerir la realización de un trabajo previo de preparación y finalizar redacción de un informe de laboratorio o la inclusión de las distintas experiencias en un cuaderno de laboratorio.

## EVALUACIÓN Y CRITERIOS DE CALIFICACIÓN

Actividades de evaluación	Criterios de evaluación	PESO
Realización de exámenes:		
Tests Al final de cada tema	Comprensión de conceptos.	
Examen Final	Aplicación de conceptos a la resolución de problemas prácticos.	40%
	Presentación y estructura.	
Evaluación del Rendimiento.	Comprensión de conceptos.	
	Aplicación de conceptos a la resolución de problemas prácticos.	
Prácticas de laboratorio.	Presentación y estructura.	60%
	Informe de las prácticas	

## Calificaciones

### Criterios de Calificación

La calificación en la convocatoria ordinaria de la asignatura se obtendrá como:

Exámenes (40% del total):



# COMILLAS

UNIVERSIDAD PONTIFICIA

ICAI

ICADE

CIHS

GUÍA DOCENTE

2024 - 2025

Tests a lo largo del curso de cada bloque

Es imprescindible sacar un 5 o más en esta parte para hacer media.

Evaluación del rendimiento (60% del total):

Informes de prácticas y explicación en clase

Es imprescindible sacar un 4 o más en esta parte para hacer media.

Convocatoria Extraordinaria

Examen (85% del total):

**Teoría (85%)**

**La nota mínima del examen de Teoría extraordinario es de 4.**

**Rendimiento en clase (15%)**

## BIBLIOGRAFÍA Y RECURSOS

### Bibliografía Básica

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with python*. O'Reilly Media.

Chollet, F. (2017). *Deep learning with python*. Manning Publications.

Geron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras and TensorFlow: concepts, tools, and techniques to build intelligent systems* (2nd ed.). O'Reilly.

Osinga, D. (2018). *Deep Learning Cookbook : Practical Recipes to Get Started Quickly*. O'Reilly.

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>