



GENERAL INFORMATION

Data of the subject	
Subject name	Unstructured Data Analysis
Subject code	DTC-MBD-522
Main program	Master in Big Data Technologies and Advanced Analytics
Involved programs	Máster en Big Data. Tec. y Analítica Avanzada/Master in Big Data Technologies and Advanced Analytics [First year]
Credits	6,0 ECTS
Type	Optativa
Department	Department of Telematics and Computer Sciences

Teacher Information	
Teacher	
Name	Ana Laguna Pradas
Department	Department of Telematics and Computer Sciences
E-Mail	alaguna@icai.comillas.edu
Teacher	
Name	Cristina Puente Águeda
Department	Department of Telematics and Computer Sciences
Office	Alberto Aguilera 25
E-Mail	cristina.puente@icai.comillas.edu
Phone	4268

DESCRIPTION OF THE SUBJECT

Contextualization of the subject

Course contents

Contents
LOGS
Intro to Data Logs Analysis
Acquiring Data Logs
Pre-processing Data Logs

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

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

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

GRAPHS

Intro to Network Analysis and Graph theory

EVALUATION AND CRITERIA

Grading

Assessment Activities	Criteria	Weight
-----------------------	----------	--------

Realization of exams: Tests

- At the end of each topic
- Final Exam
 - Understanding of concepts.
 - Application of concepts to the resolution of practical problems.
 - Presentation and structure.

Weight: 40%

Laboratory practices

- Understanding of concepts.
- Application of concepts to the resolution of practical problems.
- Presentation and structure.
- Laboratory Practices Report

Weight: 60%

BIBLIOGRAPHY AND RESOURCES

In compliance with current regulations on the **protection of personal data**, we would like to inform you that you may consult the aspects related to privacy and data [that you have accepted on your registration form](#) by entering this website and clicking on "download"

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