

Curso Académico: ( 2023 / 2024 )

Fecha de revisión: 28-04-2023

Departamento asignado a la asignatura: Departamento de Teoría de la Señal y Comunicaciones

Coordinador/a: PELAEZ MORENO, CARMEN

Tipo: Optativa Créditos ECTS : 6.0

Curso : 4 Cuatrimestre :

## OBJETIVOS

Los datos audiovisuales son de vital importancia para la industria del entretenimiento (medios digitales, televisión, radio, podcasts, videojuegos, música, etc.) en las que su combinación con las telecomunicaciones ha cambiado de forma radical nuestras vidas, especialmente en la manera en la que interactuamos con dichos datos (asistentes inteligentes como Siri, Alexa, Google Assistant, etc).

Además, estos datos son cada vez más relevantes en áreas como la medicina en donde la sofisticación cada vez mayor de dispositivos de sensado e incluso de dispositivos wearable tales como relojes inteligentes, gafas de realidad virtual y aumentada, etc., generan cada vez más tipos de datos con una inmensa potencialidad de transformación de la sociedad y de creación de nuevos mercados.

Pero la tecnología clave que se ha hecho imprescindible para transformar esta gran cantidad de datos audiovisuales en información y conocimiento útiles es la inteligencia artificial o el aprendizaje máquina, incluyendo redes neuronales y aprendizaje profundo.

Por ello, el objetivo de esta asignatura es proporcionar a los estudiantes el conocimiento teórico y metodológico sobre algoritmos y métodos para el análisis de la información audiovisual incluyendo recuperación e indexado de información multimedia para navegación y búsqueda, perfilado de usuarios, minería de opinión y posicionamiento, personalización de recomendaciones, etc.

Además, se adoptará un punto de vista eminentemente práctico proporcionando las herramientas para poner en práctica los conocimientos teóricos en el laboratorio, de manera que acaben siendo capaces de desarrollar un proyecto de análisis de datos audiovisuales basado en el aprendizaje máquina.

Finalmente, esto permitirá establecer conexiones con la gran cantidad de aplicaciones que poseen y los productos y servicios comerciales que soportan (por ejemplo, varios servicios de Google, plataformas como Twitter, Instagram, TikTok, Spotify, Netflix, YouTube, Twitch, Shazam y un largo etc.).

Tras la finalización satisfactoria del curso, el alumnado habrá adquirido (o progresado en la adquisición -en el caso de las transversales-) las siguientes competencias:

### 1. COMPETENCIAS TRANVERSALES/GENÉRICAS:

- 1.1. Capacidad de trabajo personal.
- 1.2. Capacidad de análisis y síntesis.
- 1.3. Capacidad para aplicar conceptos teóricos en casos prácticos.
- 1.4. Destrezas relacionadas con el trabajo en grupo y la colaboración con otros compañeros.
- 1.5. Destrezas relacionadas con la realización de presentaciones orales y escritas.

### 2. OBJETIVOS ESPECÍFICOS:

- 2.1. Comprender los fundamentos del análisis de datos audio-visuales y sus aplicaciones.
- 2.2. Comprender los métodos básicos de representación y descripción de la voz, el audio, la imagen y el vídeo.
- 2.3. Comprender los métodos y tecnologías empleados para clasificación, detección, indexado, recuperación, filtrado, personalización o reconocimiento de voz, audio, imagen o vídeo.
- 2.4. Capacidad para diseñar e implementar los métodos y tecnologías anteriores en problemas

prácticos de análisis automático de voz, audio, imagen y vídeo.

## DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

El problema moderno de la sobrecarga de información provocado por la disponibilidad de cantidades inabarcables de información a través de internet hace que sean necesarios sistemas que nos permitan encontrar la información que buscamos y filtrarla o personalizarla en función de nuestras necesidades. Para ello es fundamental poder indexar automáticamente contenidos no sólo textuales sino también de audio, voz, imagen o vídeo, recurriendo a métodos basados en el propio contenido o en etiquetado colaborativo como el que tiene lugar en las redes sociales.

- Tema 0. Panorámica del análisis de datos audiovisuales.
- Tema 1. Descriptores de datos audiovisuales
- Tema 2. Métodos para el análisis de datos audiovisuales
- Tema 3. Sistemas de recuperación y filtrado de datos audiovisuales
- Tema 4. Aplicaciones

## ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS

Se proponen varios tipos de actividades formativas: clases de teórico-prácticas, prácticas guiadas en laboratorio y proyecto final de laboratorio.

Se adoptarán varios tipos de metodologías: clases magistrales, aprendizaje invertido y aprendizaje basado en problemas (con cantidad variable de supervisión y alcance).

### CLASES DE TEORICO-PRACTICAS (3 ECTS)

Las clases de teoría proporcionan una visión general de los principales conceptos teóricos y matemáticos en la representación y procesado de datos audiovisuales así como de las herramientas analíticas utilizadas para el indexado y la recuperación así como para el perfilado de usuarios y la recomendación.

En estas clases se introducirán ejemplos en el laboratorio dentro de las exposiciones teóricas: todas las sesiones formativas que sea posible se van a llevar a cabo en el laboratorio lo que nos permite que todos los alumnos puedan utilizar un ordenador. Los ejemplos estarán imbricados en la exposición teórica proporcionando gran dinamismo a la clase. Por otra parte, esto nos permite introducir conceptos y herramientas básicas que solucionen las diferencias en formación previa.

Además, todos los temas se iniciarán con un debate sobre sus implicaciones tecnológicas. Para hacerlo efectivo se utilizarán técnicas de aprendizaje invertido (flipping the classroom). En concreto, se proporcionará previamente al alumnado un vídeo seleccionado para provocar el debate y la participación, junto con una serie de preguntas que les ayuden a reflexionar y que el profesorado no responderá de forma concluyente. De esta manera se pretende provocar curiosidad por averiguar los componentes de las tecnologías que se explicarán en las siguientes sesiones de cada tema.

### APRENDIZAJE COLABORATIVO (1 ECTS)

Se construirá un ejemplo práctico de la utilización de los métodos colaborativos aprendidos en las clases de teoría para el etiquetado de contenidos, para establecer conexiones entre las aplicaciones y servicios comerciales que se benefician del análisis de datos audiovisuales, así como de los hitos y fracasos de las tecnologías aprendidas. Para ello se utilizarán herramientas colaborativas de mapas conceptuales.

### PROYECTO FINAL (2 ECTS)

Los alumnos realizarán un proyecto en el que programarán un sistema modular completo relacionado con las temáticas explicadas en clase. Se proporcionará a los alumnos unas guías básicas y se impartirán unas sesiones preparatorias utilizando las técnicas del Problem Based Learning.

## SISTEMA DE EVALUACIÓN

|  |     |
|--|-----|
| <b>Peso porcentual del Examen Final:</b>           | 0   |
| <b>Peso porcentual del resto de la evaluación:</b> | 100 |
| Prácticas y cuestionarios (40%)                    |     |
| Proyecto final (60%)                               |     |

## BIBLIOGRAFÍA BÁSICA

- C. D. Manning, P. Raghavan and H. Schultze Introduction to Information Retrieval, MIT press, 2008
- Rafael C. González and Richard E. Woods Digital Image Processing, Fourth Edition, Pearson, 2018

#### BIBLIOGRAFÍA COMPLEMENTARIA

- Ben Gold (Author), Nelson Morgan (Author), Dan Ellis (Author) Speech and Audio Signal Processing: Processing and Perception of Speech and Music, Wiley, 2011
- Dan Jurafsky and James H. Martin Speech and Language Processing (3rd ed.), Prentice Hall, 2018
- Li Deng (Editor), Yang Liu (Editor) Deep Learning in Natural Language Processing, Springer, 2018
- Ricardo Baeza-Yates, Berthier Ribeiro-Neto Modern Information Retrieval: the concepts and technology behind search, 2nd Edition, Pearson, 2011
- S. Theodoridis and K. Koutroumbas Pattern Recognition, 4th ed., Academic Press, 2008
- Wilhelm Burger and Mark J. Burge Principles of Digital Image Processing: Fundamental Techniques, Springer-Verlag, 2009
- Yu, Dong, Deng, Li Automatic Speech Recognition, Springer, 2015