

Curso Académico: (2020 / 2021)

Fecha de revisión: 09-07-2020

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

Coordinador/a: DIAZ DE MARIA, FERNANDO

Tipo: Optativa Créditos ECTS : 6.0

Curso : 4 Cuatrimestre :

OBJETIVOS

El objetivo del curso es proporcionar a los estudiantes el conocimiento teórico y metodológico sobre algoritmos y métodos para recuperación e indexado de información multimedia.

Tras la finalización satisfactoria del curso, el alumno 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 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.

CB1, CB2

CG3, CG11

ETEGITT9, ETEGITT3

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 (música, voz, etc.), imagen o vídeo.

El objetivo de esta asignatura es introducir las técnicas básicas de tratamiento de voz, audio, imagen y vídeo, con una marcada orientación práctica; para ello se apoyará en una metodología de aprendizaje basado en proyectos. En particular, se presentarán, tanto en el aula como el laboratorio, los métodos necesarios para llevar a cabo dos proyectos, uno en el ámbito de tratamiento de imagen y otro en el ámbito de tratamiento de audio:

- Imagen: reconocimiento de caras, construcción de imágenes panorámicas, detección de vehículos, etc.
- Audio: sistemas de clustering o clasificación de audio por géneros, clasificación de electrocardiogramas, clasificación de emociones, etc.

Ambos proyectos se plantearán en la plataforma Kaggle, a modo de *¿challenges¿*, de modo que los estudiantes puedan competir entre sí.

La asignatura se cerrará con una lección introductoria a las redes neuronales y sus aplicaciones en

tratamiento de voz, audio, imagen y vídeo, que tendrá su continuación en dos asignaturas optativas del segundo cuatrimestre:

- Aprendizaje profundo para el análisis de imágenes
- Procesamiento del lenguaje natural

Programa de la asignatura

1. Introducción al análisis de datos audiovisuales (audio & visual analytics)
2. Datos audiovisuales: representación digital
3. Tratamiento digital de imagen y vídeo
 - 3.1. Operaciones punto a punto y filtros
 - 3.2. Segmentación de imágenes y procesado morfológico
 - 3.3. Extracción de características
 - 3.4. Proyecto integrador (por ejemplo: reconocimiento de caras, construcción de imágenes panorámicas, detección de vehículos, etc.)
4. Tratamiento de voz y audio
 - 4.1. Producción de habla y percepción de audio
 - 4.2. Análisis localizado en el tiempo. ¿Cómo funciona Shazam?
 - 4.3. Extracción de características
 - 4.4. Proyecto integrador: (p. ej., clustering o clasificación de audio, clasificación de electrocardiogramas, clasificación de emociones, etc.)
5. Introducción a las Redes Neuronales para análisis de voz, audio, imagen y vídeo

ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS

Se proponen varios tipos de actividades formativas: clases de teoría y problemas, prácticas en laboratorio y proyecto final de laboratorio.

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

CLASES DE TEORÍA (2.5 ECTS)

Las clases de teoría proporcionan una visión general de los principales conceptos teóricos y matemáticos así como de las herramientas básicas de análisis de voz, audio, imagen y vídeo

PRÁCTICAS GUIADAS DE LABORATORIO (1.75 ECTS)

Las prácticas de laboratorio se han diseñado con el propósito de que los alumnos apliquen las herramientas matemáticas presentadas en las clases teóricas a casos prácticos. Los estudiantes aprenderán a utilizar diferentes métodos de análisis de audio e imagen, como clustering de audio, reconocimiento de caras o indexado de texto, y a interpretar los resultados que obtengan en las prácticas.

PROYECTOS (1.75 ECTS)

Los alumnos realizarán por grupos un sistema de análisis imagen y otro de análisis de audio .

SISTEMA DE EVALUACIÓN

En consonancia con la orientación a proyecto de la asignatura, ésta se evaluará exclusivamente en función de los dos proyectos realizados:

50% Proyecto 1: memoria, resultados y presentación del mismo.

50% Proyecto 2: memoria, resultados y presentación del mismo.

Peso porcentual del Examen Final:	0
Peso porcentual del resto de la evaluación:	100

BIBLIOGRAFÍA BÁSICA

- C. D. Manning, P. Raghavan and H. Schultze Introduction to Information Retrieval, MIT press, 2008
- N. Morgan and B. Gold Speech and Audio Signal Processing: Processing and Perception of Speech and Music, John Wiley & Sons, Inc. New York, NY, USA, 1999
- Rafael C. González and Richard E. Woods Digital Image Processing, Fourth Edition, Pearson, 2018

BIBLIOGRAFÍA COMPLEMENTARIA

- Ricardo Baeza-Yates, Berthier Ribeiro-Neto Modern Information Retrieval: the concepts and technology behind search, 2nd Edition, Pearson, 2011
- S. Theodoridis and K. Koutroubas Pattern Recognition, 4th ed., Academic Press, 2008

