

**FICHA IDENTIFICATIVA****Datos de la Asignatura**

<b>Código</b>	34907
<b>Nombre</b>	Procesado digital de audio y voz
<b>Ciclo</b>	Grado
<b>Créditos ECTS</b>	6.0
<b>Curso académico</b>	2023 - 2024

**Titulación(es)**

<b>Titulación</b>	<b>Centro</b>	<b>Curso</b>	<b>Periodo</b>
1403 - Grado de Ingeniería Telemática	Escuela Técnica Superior de Ingeniería	4	Segundo cuatrimestre

**Materias**

<b>Titulación</b>	<b>Materia</b>	<b>Caracter</b>
1403 - Grado de Ingeniería Telemática	19 - Optatividad	Optativa

**Coordinación**

<b>Nombre</b>	<b>Departamento</b>
COBOS SERRANO, MAXIMO	240 - Informática

**RESUMEN**

La asignatura “Procesado Digital de Audio y Voz” es una asignatura de cuarto curso que forma parte de la oferta en optatividad del Grado en Ingeniería Telemática. La asignatura complementa los contenidos vistos en otras asignaturas del Grado como “Procesado Digital de la Señal”, “Señales y Sistemas Lineales” y “Fundamentos Matemáticos de las Comunicaciones”, ofreciendo una visión aplicada de los conceptos estudiados a lo largo de estas materias. Así, los temas abarcados por esta asignatura están orientados a la aplicación del tratamiento digital de señales y el aprendizaje automático en el campo del procesamiento de audio y de voz.

La asignatura justifica la importancia del procesamiento digital de señales de audio en los sistemas multimedia actuales, revisando brevemente algunos conceptos básicos estudiados en cursos anteriores. La estructuración de los contenidos sigue un enfoque en el que se comienza introduciendo los conceptos más teóricos en la primera parte del curso relacionados con el procesamiento digital de señales, dando paso a otros temas más enfocados al aprendizaje máquina, los cuales han adquirido una gran importancia en la actualidad. Así, se parte de una revisión de conceptos básicos de procesamiento de señal, introduciendo varias representaciones de las señales de audio, poniendo especial atención a las representaciones tiempo-



frecuencia. La parte teórica se complementa con el estudio del sistema auditivo humano y su impacto en el diseño de sistemas prácticos de codificación de audio con pérdidas, los fundamentos del sistema fonador y los modelos fuente/filtro utilizados en codificación de voz.

La segunda parte del curso introduce los fundamentos de los sistemas de audio basados en aprendizaje máquina, introduciendo el uso de características utilizadas ampliamente en problemas de audio y su combinación con clasificadores binarios clásicos como las máquinas de vectores soporte o regresión logística. Por último se introduce al alumno a los sistemas de aprendizaje profundo actuales.

## CONOCIMIENTOS PREVIOS

### Relación con otras asignaturas de la misma titulación

No se han especificado restricciones de matrícula con otras asignaturas del plan de estudios.

### Otros tipos de requisitos

## COMPETENCIAS

## RESULTADOS DE APRENDIZAJE

Esta asignatura permite profundizar en los siguientes resultados fundamentales de aprendizaje . Nótese, que dado que se trata de una asignatura optativa, no existen competencias específicas asociadas a las mismas.

-Adquirir nuevos conocimientos y técnicas adecuados para el diseño, el desarrollo y la implementación de sistemas de telecomunicación, en especial, aquellos relacionados con las comunicaciones multimedia.

-Seguir el proceso de diseño de sistemas de procesado basados en percepción auditiva y visual, relacionando los aspectos envueltos en el mismo con la teoría matemática asociada.

Como complemento a los resultados anteriores, esta asignatura también permite adquirir las siguientes destrezas y habilidades sociales:

-Aplicar el conocimiento sobre percepción auditiva y visual en el diseño de sistemas de procesado de señal.

-Identificar las limitaciones en la conversión analógica/digital de señales de audio y voz.

-Seleccionar adecuadamente los parámetros necesarios para la codificación y almacenamiento de señales de audio.



- Identificar los modelos de los procesos que guían el diseño de los actuales sistemas de telecomunicación.
- Implementar sistemas básicos de procesado de señal orientados al tratamiento de audio.
- Describir adecuadamente los principios que rigen la compresión de señales de audio, relacionándolos con la teoría matemática y física asociada.
- Fomentar el trabajo en equipo y de la organización en tareas y subtareas.

## DESCRIPCIÓN DE CONTENIDOS

### 0. Introducción

Visión global del procesado digital de señal de audio, imágenes y vídeo. El procesado de señal en sistemas multimedia.

### 1. Revisión de procesado de señal

Descripción de contenidos (Castellano):  
Introducción. Revisión de convolución y filtrado. Teoría del muestreo. Procesado de señal en tiempo discreto. DFT y DTFT. Procesos aleatorios.

### 2. Audio y principios psicoacústicos

Descripción de contenidos (Castellano):  
Introducción. Fundamentos de acústica. Audición y percepción. Sonoridad y bandas críticas.

### 3. Introducción al aprendizaje automático en audio

Procesado de señal e inteligencia artificial en audio. Evolución e historia de los sistemas de audio basados en aprendizaje automático. Retos relacionados con el uso de las técnicas de aprendizaje automático en audio. Aplicaciones actuales.

### 4. Representaciones de audio para aprendizaje automático

Representaciones de audio tiempo-frecuencia: STFT, Espectrograma Mel, Bancos de filtros Gammatone, Chromagrama. Características clásicas: temporales, espectrales, MFCCs. Representaciones compactas y reducción de dimensionalidad: NMF, PCA.

**5. Técnicas de aprendizaje automático aplicadas a audio y voz**

Introducción al aprendizaje supervisado y no supervisado. Clasificadores binarios lineales: SVM, Regresión logística. Clasificadores no lineales: k-NN, modelos Kernel. Modelos de mezclas de Gaussianas (GMM), k-Means. Introducción a las redes neuronales y el aprendizaje profundo.

**VOLUMEN DE TRABAJO**

ACTIVIDAD	Horas	% Presencial
Clases de teoría	30,00	100
Prácticas en laboratorio	20,00	100
Prácticas en aula	10,00	100
Elaboración de trabajos en grupo	20,00	0
Elaboración de trabajos individuales	10,00	0
Preparación de actividades de evaluación	20,00	0
Preparación de clases de teoría	10,00	0
Preparación de clases prácticas y de problemas	15,00	0
Resolución de casos prácticos	15,00	0
<b>TOTAL</b>	<b>150,00</b>	

**METODOLOGÍA DOCENTE**

1.- Trabajo presencial formado por:

- Clases de teoría, las cuáles consistirán en la presentación y explicación básica de la materia correspondiente. Se propondrán actividades de corta duración, las cuales exigirán la intervención de los alumnos con el objetivo de confirmar la comprensión de la teoría expuesta.
- Clases de ejercicios y seminarios prácticos, diseñadas para resolver problemas de mayor envergadura o bien temporal o bien conceptual.
- Clases de laboratorio, pensadas para comprobar experimentalmente algunas de las cuestiones más relevantes vistas en las clases de teoría.



2.- Trabajo no presencial formado por:

a) Preparación de cuestionarios de tema a realizar tras finalizar cada unidad didáctica.

b) Preparación de las prácticas de laboratorio, para las que el alumno deberá haber leído y asimilado el contenido del boletín de prácticas, así como haber repasado la teoría relevante.

Tutorías individuales y/o colectivas. Se establecen unas determinadas horas de tutorías no programadas por semana a las que los alumnos podrán asistir para aclarar sus dudas (presencialmente o en modo on-line).

## EVALUACIÓN

El mecanismo de evaluación está basado en una evaluación continua del trabajo desarrollado por el alumno a lo largo del curso. Se tienen en cuenta los siguientes ítems y valoraciones:

a)- Resultado de cuestionarios de tema realizados tras la finalización de cada unidad temática (50% de la nota final)

b)- Realización de informes de prácticas (25% de la nota final)

c)- Realización de cuestionario final global (25% de la nota final)

Para superar la asignatura se deberá tener una nota media en los ítems anteriores superior a 5. Los mínimos requeridos para superar la asignatura serán el equivalente a un 5 sobre 10 tanto en la nota media obtenida en los cuestionarios de tema (a), como en el cuestionario final (c).

La nota de prácticas (b) no será recuperable en segunda convocatoria.



## REFERENCIAS

### Básicas

- Virtanen, Plumbley, Ellis, (eds): Computational Analysis of Sound Scenes and Events, Springer, 2018.
- Giannakopoulos, T. And Pikrakis, A. Introduction to Audio Analysis, Elsevier, 2014.
- Zölzer, Udo., Digital Audio Signal Processing, 2nd edition, Wiley, 2008. ISBN: 0470997850
- Pulkki, V., Karjalainen, M. Communication Acoustics: An Introduction to Speech, Audio and Psychoacoustics. Wiley (2015). ISBN: 978-1-118-86654-2

### Complementarias

- Aurélien Géron, Hands-on Machine Learning with Scikit-Learn, Keras & Tensorflow, OReilly, 2020.
- -Smith III, Julius O., Spectral Audio Signal Processing, W3K Publishing, 2011. 978-0974560731
- -Bosi, M. and Goldberg, Richard E., Introduction to Digital Audio Coding and Standards, Kluwer Academic Publishers, 2003. ISBN: 978-1402073571