

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

Código	36439
Nombre	Análisis y procesamiento de audio y voz
Ciclo	Grado
Créditos ECTS	4.5
Curso académico	2021 - 2022

Titulación(es)

Titulación	Centro	Curso	Periodo
1406 - Grado en Ciencia de Datos	Escuela Técnica Superior de Ingeniería	4	Segundo cuatrimestre

Materias

Titulación	Materia	Caracter
1406 - Grado en Ciencia de Datos	16 - Análisis y Procesamiento de Audio y Voz	Optativa

Coordinación

Nombre	Departamento
ROMERO GOMEZ, VERONICA	240 - Informática
SEGURA GARCIA, JAUME	240 - Informática

RESUMEN

La asignatura “Análisis y procesamiento de Audio y Voz” es una asignatura de cuarto curso que forma parte de la oferta en optatividad del Grado en Ciencia de Datos. La asignatura complementa los contenidos vistos en otras asignaturas del Grado como, “Señales y Sistemas”, “Aprendizaje máquina” y “Procesado del lenguaje natural”, 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 en el campo del procesado de audio y de voz. La asignatura justifica la importancia del procesado digital de señales de audio en los sistemas multimedia y de procesado masivo de datos 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, dando paso a temas que utilizan tales conceptos en sistemas prácticos actuales. Así, se parte de una revisión de conceptos básicos de procesado de señal, así como otros conceptos relevantes en audio como el muestreo, la cuantificación, las técnicas de análisis/síntesis en tiempo corto y métodos de procesado estadístico. Se continua 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. Además, se introducen Conceptos como filtrado óptimo, predicción lineal o análisis sub-banda son introducidos desde un punto de vista aplicado. La parte teórica se complementa con los fundamentos de los sistemas de reconocimiento de voz basados en modelos ocultos de Markov y modelos de lenguaje. Además, se estudian técnicas de aprendizaje máquina con aplicación al reconocimiento de locutores y síntesis de voz. El objetivo de la asignatura es proporcionar al alumno una visión aplicada de los conceptos estudiados a lo largo del Grado, facilitando la comprensión de términos abstractos mediante el estudio de aplicaciones concretas. Así, se pretende que el alumno obtenga una visión práctica del estudio del procesado y caracterización de la señal voz, modelos de lenguaje, aprendizaje máquina, reconocimiento y síntesis de voz.

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

1406 - Grado en Ciencia de Datos

- (CG01) Conocimiento de materias básicas y tecnologías, que le capacite para el aprendizaje de nuevos métodos y tecnologías, así como que le dote de una gran versatilidad para adaptarse a nuevas situaciones.
- (CT01) Ser capaces de acceder a herramientas de información (bibliográficas) y de utilizarlas apropiadamente en el desarrollo de sus tareas cotidianas.
- (CE03) Capacidad para resolver problemas de clasificación, modelización, segmentación y predicción a partir de un conjunto de datos.
- (CE10) Capacidad para procesar señales de forma digital extrayendo información de ellas.
- (CB5) Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía.

RESULTADOS DE APRENDIZAJE

Aplicar técnicas de estimación y detección sobre señales de audio (CB05, CE10), entendiendo los fundamentos estadísticos asociados a las mismas. Conocer los fundamentos de los sistemas de reconocimiento de voz basados en modelos ocultos de Markov. Aplicar técnicas de aprendizaje máquina a señales de audio y voz, con aplicación al reconocimiento de locutores y otros problemas de clasificación (CB05, CE03). Comprender los fundamentos básicos de la generación y la percepción de señales de audio en general, voz y música (CB05, CE10). Conocer la representación de señales de audio en el dominio temporal y frecuencial, así como los fundamentos matemáticos asociados a ambos dominios. Aplicar transformaciones tiempo-frecuencia para el análisis en tiempo corto de series temporales (CB05, CE10).



Aplicar técnicas de estimación y detección sobre señales de audio, entendiendo los fundamentos estadísticos asociados a las mismas (CE03, CE10).

DESCRIPCIÓN DE CONTENIDOS

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

Análisis de la señal de voz. Reconocimiento automático del habla. Modelado acústico y del lenguaje. Reconocimiento del locutor Síntesis del habla. Representaciones vectoriales de las palabras. Descriptores para señales de audio y musicales. Análisis de señales de audio. Información y recuperación en bases de datos de audio. Segmentación y detección de eventos en tramas de audio.

2. Revisión de procesado de señal

Introducción. Revisión de convolución y filtrado. Teoría del muestreo. Procesado de señal en tiempo discreto. DFT y DTFT. Procesos aleatorios.

3. Análisis y síntesis de audio.

Introducción. Short-Time Fourier Transform. Bancos de filtros. Overlap-add.

4. Codificación.

Introducción. Predicción lineal. Codificación subbanda y transformadas. MPEG Audio

5. Reconocimiento de voz.

Introducción. Aproximación estadística. Modelos Ocultos de Markov. Modelado de lenguaje. Decodificación. Introducción a las redes neuronales.

6. Reconocimiento del locutor.

Introducción. Identificación del Locutor. Verificación del Locutor. Dependiente del texto e independiente del texto.

7. Síntesis de voz.

Introducción. Preproceso. Procesado del lenguaje natural. Síntesis estadística.

**VOLUMEN DE TRABAJO**

ACTIVIDAD	Horas	% Presencial
Clases de teoría	22,00	100
Prácticas en laboratorio	15,00	100
Prácticas en aula	8,00	100
Asistencia a eventos y actividades externas	2,00	0
Elaboración de trabajos en grupo	6,00	0
Elaboración de trabajos individuales	36,50	0
Preparación de actividades de evaluación	2,00	0
Preparación de clases prácticas y de problemas	6,00	0
Resolución de casos prácticos	15,00	0
TOTAL	112,50	

METODOLOGÍA DOCENTE

MD1 - Actividades teóricas. Desarrollo expositivo de la materia con la participación del estudiante en la resolución de cuestiones puntuales. Realización de cuestionarios individuales de evaluación.

MD2 - Actividades prácticas. Aprendizaje mediante resolución de problemas, ejercicios y casos de estudio a través de los cuales se adquieren competencias sobre los diferentes aspectos de la materia.

MD4 - Trabajos en laboratorio y/o aula ordenador. Aprendizaje mediante la realización de actividades desarrolladas de forma individual o en grupos reducidos y llevadas a cabo en laboratorios y/o aulas de ordenador

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, 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:

- Resolución y presentación de ejercicios. Se trata de resolver los boletines de ejercicios propuestos por el profesor y/o la exposición en público de la resolución de algunos de ellos.
- Preparación de los exámenes.
- 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 duda



EVALUACIÓN

SE1 - Prueba objetiva, consistente en uno o varios exámenes que constan tanto de cuestiones teórico-prácticas como de problemas. 20%+40%

SE2 - Evaluación continua de cada alumno de las actividades prácticas a partir de la elaboración de trabajos/memorias y/o exposiciones orales. 35%

SE3 - Evaluación basada en la participación y grado de implicación del alumno en el proceso de enseñanza-aprendizaje, teniendo en cuenta la asistencia regular a las actividades presenciales previstas y la resolución de cuestiones y problemas propuestos periódicamente. 5%

El mecanismo de evaluación es el que se podría denominar tradicional adaptado, que no llega a ser una evaluación continua completa. Se tienen en cuenta los siguientes ítems y valoraciones:

Valoración de la asistencia y la participación (5% de la nota final). Nota no recuperable, se conserva la nota obtenida en segunda convocatoria.

Asistencia y realización de las prácticas (20% de la nota final). Nota no recuperable, se conserva la nota obtenida en segunda convocatoria.

Resolución de ejercicios propuestos (15% de la nota final). Nota no recuperable, se conserva la nota obtenida en segunda convocatoria.

Resultado del examen parcial (20% de la nota final).

Examen/Trabajo final (40% de la nota final).

Se considera que un alumno asiste regularmente a clase cuando no haya faltado a más del 25% de las clases dadas hasta la mitad del tiempo de impartición de la asignatura.

Para los alumnos que no puedan asistir regularmente a clase, se ofrece un modelo alternativo en el que la valoración de la asistencia y participación se sustituirán por trabajos adicionales con un porcentaje total equivalente. Los alumnos que quieran acogerse a esta opción deberán comunicarlo al profesor durante las 3 primeras semanas del curso.

El último ítem evaluable (Examen/Trabajo final) se refiere a la posibilidad de que el alumno elija entre presentarse a un examen final o bien realice un trabajo propuesto por el profesor. El trabajo final incluye la programación de algoritmos de procesado de señal, un informe técnico y una presentación oral, así como un pequeño examen relacionado con el tema escogido para el trabajo.

Los mínimos requeridos para superar la asignatura serán el equivalente a un 3.5 sobre 10 tanto en el examen parcial como en el trabajo/examen final. El resto de ítems evaluables no están sometidos a mínimo.

Tanto en primera como en segunda convocatoria, aquellos estudiantes que no superen el 5 mediante las actividades de evaluación continua, podrán elegir entre dos opciones:

- Examen final+parcial (40%/20%). Se guarda la nota del examen parcial o del trabajo en caso de ser superior a 3.5 y se recupera la parte suspendida mediante examen.

- Examen final (60%). La nota correspondiente al examen parcial y a la del trabajo/examen se recupera con un único examen final.

En cualquier caso, el sistema de evaluación se regirá por lo establecido en el Reglamento de Evaluación y Calificación de la Universidad de Valencia para Grados y Másteres

(<https://webges.uv.es/uvTaeWeb/MuestraInformacionEdictoPublicoFrontAction.do?accion=inicio&idEdictoSeleccionado=5639>)



REFERENCIAS

Básicas

- Wendy Holmes. Speech synthesis and recognition [2nd ed]
- Sadaoki Furui. Digital Speech Processing, Synthesis and Recognition [2 ed.]
- Paul Hill. Audio and speech processing with MATLAB [1st edition]

ADENDA COVID-19

Esta adenda solo se activará si la situación sanitaria lo requiere y previo acuerdo del Consejo de Gobierno

La metodología docente de la asignatura seguirá el Modelo Docente aprobado por la Comisión Académica de Título de Ciencia de Datos (<https://go.uv.es/cienciadatos/ModelDocentGCD>).

En caso de que se produzca un cierre de las instalaciones por causas sanitarias que afecte total o parcialmente a las clases de la asignatura, estas serán sustituidas por sesiones no presenciales siguiendo los horarios establecidos.

Si el cierre afectara a alguna prueba de evaluación presencial de la asignatura, esta será sustituida por una prueba de naturaleza similar que se realizará en modalidad virtual a través de las herramientas informáticas soportadas por la Universitat de València. Los porcentajes de cada prueba de evaluación permanecerán invariables, según lo establecido por esta guía.