



Guía docente

320123 - TP - Tecnología del Habla

Última modificación: 15/07/2024

Unidad responsable: Escuela Superior de Ingenierías Industrial, Aeroespacial y Audiovisual de Terrassa
Unidad que imparte: 739 - TSC - Departamento de Teoría de la Señal y Comunicaciones.

Titulación: GRADO EN INGENIERÍA DE SISTEMAS AUDIOVISUALES (Plan 2009). (Asignatura optativa).

Curso: 2024 **Créditos ECTS:** 6.0 **Idiomas:** Catalán, Castellano

PROFESORADO

Profesorado responsable: Ignasi Esquerra

Otros: Albino Nogueiras

CAPACIDADES PREVIAS

Es conveniente haber cursado las asignaturas de "Procesado digital de audio" y "Algorísmica y programación audiovisual" y tener cierto conocimiento de programación en entornos Linux.

COMPETENCIAS DE LA TITULACIÓN A LAS QUE CONTRIBUYE LA ASIGNATURA

Específicas:

CE26. AUD: Conocimientos y capacidades para profundizar en tecnologías específicas del ámbito.

METODOLOGÍAS DOCENTES

Aprendizaje dirigido:

a) Sesiones en el laboratorio. Presentación de los contenidos teóricos, y realización de demostraciones. Ejercicios prácticos. Ejercicios de evaluación.

Aprendizaje autónomo:

b) Estudio individual, resolución de ejercicios y preparación de los trabajos.

OBJETIVOS DE APRENDIZAJE DE LA ASIGNATURA

Conocer las tecnologías utilizadas en los sistemas de interacción con los ordenadores a través de la voz. Los estudiantes en acabar el curso conocerán diversas técnicas de procesado y serán capaces de aplicarlos en sistemas de reconocimiento de voz y síntesis de voz.

HORAS TOTALES DE DEDICACIÓN DEL ESTUDIANTADO

| Tipo | Horas | Porcentaje |
|----------------------------|-------|------------|
| Horas grupo pequeño | 30,0 | 20.00 |
| Horas aprendizaje autónomo | 90,0 | 60.00 |
| Horas grupo grande | 30,0 | 20.00 |

Dedicación total: 150 h

CONTENIDOS

1. INTERACCIÓN ORAL CON LOS ORDENADORES

Descripción:

Tecnologías del habla y del lenguaje. Perspectiva histórica de la interacción persona-ordenador. Aplicaciones y líneas de investigación actuales.

Dedicación: 10h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Aprendizaje autónomo: 6h

2. BASES DE DATOS

Descripción:

Bases de datos de texto y señales de voz. Fonética y lingüística. Producción y percepción del habla. Representación y análisis de las señales de voz.

Dedicación: 30h

Grupo grande/Teoría: 8h

Grupo pequeño/Laboratorio: 8h

Aprendizaje autónomo: 14h

3. SÍNTESIS DE VOZ

Descripción:

Análisis del texto. Transcripción fonética. Prosodia. Síntesis por selección de unidades. Síntesis por modelos probabilísticos.

Dedicación: 46h

Grupo grande/Teoría: 8h

Grupo pequeño/Laboratorio: 8h

Aprendizaje autónomo: 30h

4. RECONOCIMIENTO DE VOZ

Descripción:

Parametrización. Modelos acústicos. Modelos ocultos de Markov. Modelos de lenguaje.

Dedicación: 64h

Grupo grande/Teoría: 12h

Grupo pequeño/Laboratorio: 12h

Aprendizaje autónomo: 40h

SISTEMA DE CALIFICACIÓN

La calificación final es la suma ponderada de las calificaciones parciales de los temas. Cada tema es evaluado con varios trabajos de laboratorio o controles. Ningún elemento de calificación tiene un peso superior a 25% sobre el total.

Tema 1 (10%), Tema 2 (20%), Tema 3 (35%), Tema 4 (35%)



BIBLIOGRAFÍA

Complementaria:

- Huang, X.; Acero, A.; Hon, H-W. Spoken language processing: a guide to theory, algorithm and system development. Upper Saddle River: Prentice Hall, 2001. ISBN 0130226165.
- O'Shaughnessy, D. Speech communications: human and machine. 2nd ed. New York: IEEE Press, 2000. ISBN 978-0780334496.