

60936 - Tecnologías del habla

Información del Plan Docente

Año académico	2016/17
Centro académico	110 - Escuela de Ingeniería y Arquitectura
Titulación	533 - Máster Universitario en Ingeniería de Telecomunicación
Créditos	5.0
Curso	2
Periodo de impartición	Primer Semestre
Clase de asignatura	Optativa
Módulo	---

1. Información Básica

1.1. Recomendaciones para cursar esta asignatura

Es recomendable que el alumno que quiera cursar Tecnologías del Habla haya cursado o curse simultáneamente la asignatura Tratamiento de Señal para Comunicaciones y se recomienda cursar la asignatura de formación Optativa Reconocimiento de patrones en datos multimedia.

1.2. Actividades y fechas clave de la asignatura

La asignatura consta de un total de 5 créditos ECTS. Las actividades se dividen en clases teóricas, resolución de problemas o casos prácticos en clase, prácticas de laboratorio y la realización de trabajos tutelados relacionados con las tecnologías del habla. Las actividades tienen como objetivo facilitar la asimilación de los conceptos teóricos complementándolos con los prácticos, de forma que se adquieran los conocimientos y las habilidades básicas relacionadas con las competencias previstas en la asignatura.

Las fechas de inicio y finalización del curso y las horas concretas de impartición de la asignatura así como las fechas de realización de las prácticas de laboratorio e impartición de seminarios se harán públicas atendiendo a los horarios fijados por la Escuela. Las fechas de entrega y seguimiento de los trabajos prácticos tutorizados se darán a conocer con suficiente antelación en clase y en la página web de la asignatura en el anillo digital docente, <https://moodle.unizar.es/>.

2. Inicio

2.1. Resultados de aprendizaje que definen la asignatura

El estudiante, para superar esta asignatura, deberá demostrar los siguientes resultados...

R1: Comprende y domina los procesos de generación y percepción humana del habla.

R2: Conoce y utiliza las técnicas de procesado digital de la señal de voz.

R3: Comprende y utiliza las técnicas fundamentales de reconocimiento de patrones aplicadas al habla.

R4: Comprende y domina los conceptos de modelado acústico, modelado de lenguaje y su utilización en reconocimiento y síntesis del habla.

60936 - Tecnologías del habla

R5: Conoce y utiliza las tecnologías del habla relacionadas con el reconocimiento automático del habla, la síntesis de habla y el reconocimiento biométrico a partir de la voz.

2.2.Introducción

Breve presentación de la asignatura

La asignatura *Tecnologías del Habla* es una asignatura de formación Optativa que profundiza en el desarrollo de sistemas automáticos para la interacción entre personas y máquinas por medio del habla.

La asignatura proporciona al estudiante una sólida formación en conceptos fundamentales de las tecnologías del habla. Se abordan los procesos de generación del habla y percepción auditiva desde un punto de vista del proceso de comunicación con lenguaje hablado. Se presentan las técnicas de procesado de señal más comunes en el procesamiento del habla, así como las técnicas básicas de reconocimiento de patrones necesarias para el desarrollo de sistemas de reconocimiento automático del habla. Se analizan los procesos y modelos asociados al reconocimiento y síntesis del habla así como el reconocimiento biométrico a partir de la voz. El estudiante adquirirá la capacidad de entender, analizar y diseñar aplicaciones basadas en tecnologías del habla.

La asignatura combina tanto aspectos teóricos como prácticos, así que se complementarán los conceptos teóricos fundamentales con un conjunto de prácticas y trabajos tutorados.

La asignatura consta de 5 créditos ECTS, que se distribuyen en sesiones presenciales teóricas, clases de problemas, prácticas de laboratorio y trabajos prácticos en grupo.

3.Contexto y competencias

3.1.Objetivos

La asignatura y sus resultados previstos responden a los siguientes planteamientos y objetivos:

La asignatura Tecnologías del Habla plantea la adquisición del conocimiento y la comprensión de las diferentes tecnologías que componen los sistemas automáticos de interacción persona-máquina basados en el lenguaje oral. Los objetivos principales de la asignatura son alcanzar los resultados del aprendizaje expuestos previamente y la adquisición de competencias enumeradas en esta guía.

3.2.Contexto y sentido de la asignatura en la titulación

La asignatura Tecnologías del Habla proporciona a futuros profesionales ingenieros de telecomunicación los conceptos para comprender, analizar, evaluar y diseñar los aspectos fundamentales y bloques básicos de los sistemas de interacción persona-máquina basados en el lenguaje oral. Complementa de este modo los conceptos básicos adquiridos en el ámbito de la comunicación audiovisual y el desarrollo de interfaces en el Grado de Tecnologías y Servicios de Telecomunicación.

3.3.Competencias

Al superar la asignatura, el estudiante será más competente para...

CE1: Capacidad para aplicar métodos de la teoría de la información, la modulación adaptativa y codificación de canal, así como técnicas avanzadas de procesado digital de señal a los sistemas de comunicaciones y audiovisuales.

CE15: Capacidad para la integración de tecnologías y sistemas propios de la Ingeniería de Telecomunicación, con carácter generalista, y en contextos más amplios y multidisciplinares como por ejemplo en bioingeniería, conversión

60936 - Tecnologías del habla

fotovoltaica, nanotecnología, telemedicina.

CB6 : Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.

CB7 : Los estudiantes sabrán aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.

CB8: Los estudiantes serán capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios.

CB9: Los estudiantes sabrán comunicar sus conclusiones -y los conocimientos y razones últimas que las sustentan- a públicos especializados y no especializados de un modo claro y sin ambigüedades.

CB10 : Los estudiantes poseerán las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

CG4: Capacidad para el modelado matemático, cálculo y simulación en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación, desarrollo e innovación en todos los ámbitos relacionados con la Ingeniería de Telecomunicación y campos multidisciplinares afines.

CG11: Capacidad para saber comunicar (de forma oral y escrita) las conclusiones- y los conocimientos y razones últimas que las sustentan- a públicos especializados y no especializados de un modo claro y sin ambigüedades.

CG12: Poseer habilidades para el aprendizaje continuado, autodirigido y autónomo.

3.4.Importancia de los resultados de aprendizaje

La adquisición de las competencias y habilidades propuestas en la asignatura Tecnologías del Habla, así como la comprensión de los conceptos teóricos tratados, complementan las competencias de un Ingeniero de Telecomunicación en el ámbito de los sistemas de comunicación basados en la voz. Todo el conjunto de capacidades adquiridas en esta asignatura será de gran utilidad para su formación.

Los conceptos y técnicas desarrollados así como la formación práctica recibida en esta asignatura facilitarán la comprensión de los bloques integrantes de un sistema de interacción basado en el lenguaje oral y le proporcionará la base para profundizar en aspectos más detallados de los mismos y posibilidades de desarrollo profesional.

4.Evaluación

El estudiante deberá demostrar que ha alcanzado los resultados de aprendizaje previstos mediante las siguientes actividades de evaluación

E1: Prácticas de laboratorio

Las prácticas de laboratorio de la asignatura constituyen el **20% de la calificación final** . Su evaluación se realizará a

60936 - Tecnologías del habla

partir de los entregables posteriores a la realización de las mismas aportados por los alumnos y de la actitud y el rendimiento en el laboratorio, que será evaluado de forma continua. **Se requiere un nota mínima de 4 sobre 10 en este apartado** para superar la asignatura .

E2: Trabajos tutorizados

Los trabajos tutorizados representan el **30% de la calificación final** . En la calificación se valorará la capacidad analítica y crítica del alumno para estudiar un problema o aspectos concretos en un sistema basado en tecnologías del habla haciendo uso de las herramientas teóricas y prácticas aprendidas en la asignatura. Además se evaluará la originalidad de las soluciones, la capacidad para trabajar en grupo, la habilidad para coordinar el trabajo y de transmitir la información relevante de forma oral y escrita, ya que el trabajo realizado se presentará a través de un informe común al grupo y de una presentación oral. **Se requiere una nota mínima de 4 sobre 10 en este apartado** para superar la asignatura .

E3: Examen parcial

Durante el curso se realizará una prueba escrita de cuestiones teórico-prácticas que ponderará el **15% de la calificación final** . Su superación, nota mayor o igual a 5 sobre 10 en esta prueba, eximirá al estudiante de presentarse a esta parte del examen final. Esta prueba se repetirá dentro del examen final (E4) para los alumnos que no la hayan superado o deseen mejorar su nota.

E4: Examen final

El examen final consistirá en una prueba escrita que representa el **50% de la calificación final** . Una parte correspondiente al 15% de la nota se corresponde con la repetición del examen parcial (**E3**)

Se requiere una nota mínima de 4 sobre 10 en la nota del examen final (E4) para superar la asignatura .

E5: Calificación final de la asignatura.

La calificación final (CF) de la asignatura será el resultado mayor de las expresiones siguientes:

$$CF=0.20 \cdot E1 + 0.30 \cdot E2 + 0.15 \cdot E3 + 0.35 \cdot E4 \text{ ó}$$

$$CF=0.20 \cdot E1 + 0.30 \cdot E2 + 0.5 \cdot E4$$

con las restricciones comentadas: $E1 \geq 4$, $E2 \geq 4$, $E3 \geq 4$ y $E4 \geq 4$

Se dispondrá de una prueba global en cada una de las convocatorias establecidas a lo largo del curso. Las fechas y horarios vendrán determinados por el Centro.

5.Actividades y recursos

5.1.Presentación metodológica general

El proceso de aprendizaje que se ha diseñado para esta asignatura se basa en lo siguiente:

P1. Clases magistrales participativas

60936 - Tecnologías del habla

P2: Prácticas de aula

P3: Prácticas de laboratorio

P4: Trabajos prácticos tutorados

P5: Tutoría.

P6: Evaluación.

5.2. Actividades de aprendizaje

Clases magistrales participativas: Exposición por parte del profesor de los principales contenidos de la asignatura, combinada con la participación activa del alumnado. Esta actividad se realizará en el aula de forma presencial. Esta metodología, apoyada con el estudio individual del alumno está diseñada para proporcionar a los alumnos los fundamentos teóricos del contenido de la asignatura.

Prácticas de aula: en las que se realizan resolución de problemas y casos prácticos propuestos por el profesor de los fundamentos presentados en las clases magistrales, con posibilidad de exposición de los mismos por parte de los alumnos de forma individual o en grupos autorizada por el profesor. Esta actividad se realizará en el aula de forma presencial.

Prácticas de laboratorio : En las que los alumnos realizarán 5 sesiones de prácticas de 2 horas de duración en los Laboratorio de Señales y Sistemas 2.02 del Edificio Ada Byron. En grupos pequeños, se realizan una serie prácticas en las cuales se conocerán los bloques principales del sistemas de interacción basados en tecnologías del habla que permitan consolidar el conjunto de conceptos teóricos desarrollados a lo largo de las clases magistrales. Esta actividad se realizará en el Laboratorio de forma presencial.

Esta actividad se realizará de forma presencial en un aula informática. Comprenderá 5 sesiones de 2 horas de duración cada una de ellas. Los alumnos presentarán posteriormente un informe escrito que recogerá las principales conclusiones del trabajo realizado.

Trabajos prácticos tutorados : Realización de un trabajo práctico en grupo y tutorizado por el profesor, basado en los contenidos de la asignatura. Los trabajos tutorados se basarán en la comprensión, desarrollo e implementación de sistemas basados en tecnologías del habla propuestas por el profesor. Será realizado mediante un trabajo en grupo que incluirá la necesidad de coordinación y será evaluado mediante un informe escrito y una presentación oral.

Tutoría. Horario de atención personalizada al alumno con el objetivo de revisar y discutir los materiales y temas presentados en las clases tanto teóricas como prácticas.

Evaluación. Conjunto de pruebas escritas teórico-prácticas y presentación de informes o trabajos utilizados en la evaluación del progreso del estudiante. El detalle se encuentra en la sección correspondiente a las actividades de evaluación.

5.3. Programa

TEMA 1. Introducción a las tecnologías del habla

60936 - Tecnologías del habla

- Modelo de comunicación oral
- Generación del habla
- Comprensión del habla

TEMA 2. Fundamentos de reconocimiento de patrones

- Teoría de la decisión de Bayes
- Clasificadores
- Extracción y selección de parámetros
- Métodos de estimación no supervisada

TEMA 3. Procesado digital de la señal de voz

- Modelo de producción digital de la voz
- Análisis localizado de la voz: representación tiempo-frecuencia
- Predicción lineal
- Análisis Homomórfico
- Aplicaciones: estimación de pitch y formantes, reducción de ruido

TEMA 4. Reconocimiento Automático del Habla

- Perspectiva histórica y estado del arte
- Modelado acústico
- Modelado de lenguaje
- Algoritmos básicos de búsqueda
- Aplicaciones y toolkits

TEMA 5. Conversión texto-voz

- Perspectiva histórica y estado del arte
- Análisis lingüístico
- Síntesis de voz
- Aplicaciones y toolkits

TEMA 6. Reconocimiento biométrico por la voz

- Perspectiva histórica y estado del arte
- Verificación e identificación de locutor
- Aplicaciones y toolkits

5.4. Planificación y calendario

Calendario de sesiones presenciales y presentación de trabajos

El calendario de la asignatura, tanto de las horas presenciales, como las sesiones de laboratorio estará definido por el centro en el calendario académico del curso correspondiente.

5.5. Bibliografía y recursos recomendados

- Rabiner, Lawrence Richard. Digital processing of speech signals / Lawrence R. Rabiner, Ronald W. Schafer Upper Saddle River (New Jersey) : Prentice-Hall, cop. 1978
- Huang, Xuedong. Spoken language processing : a guide to theory, algorithm, and system development / Xuedong Huang, Alex Acero, Hsiao -Wuen Hon Upper Saddle River, New Jersey : Prentice Hall PTR, cop. 2001
- O'Shaughnessy, Douglas. Speech communications : human and machine / Douglas O'Shaughnessy . - 2nd ed. New York : IEEE Press, cop. 2000
- The HTK Book (for version 3.4) / S. Young...[et al.] . Cambridge University Engineering Department, 2009.
- Sphinx-4: A flexible open source framework for speech recognition / W. Walker...[et al.] Sun Microsystems Inc., Technical Report SML1 TR2004-0811, 2004
- Beigi, Homayoon. Fundamentals of Speaker Recognition / Beigi, Homayoon New York: Springer, 2011

60936 - Tecnologías del habla

- Taylor, Paul. Text-to-Speech Synthesis / Paul Taylor Academic press, 2009
- Jurafsky, Daniel. Speech and Language Processing: An Introduction to Natural Language Processing, Speech Recognition, and Computational Linguistics / Jurafsky, Daniel, and James H. Martin. - 2nd ed. Prentice-Hall, 2009