

230022 - PAV - Processament d'Àudio i Veu

Unitat responsable: 230 - ETSETB - Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona
Unitat que imparteix: 739 - TSC - Departament de Teoria del Senyal i Comunicacions
Curs: 2018
Titulació: GRAU EN ENGINYERIA DE SISTEMES AUDIOVISUALS (Pla 2009). (Unitat docent Obligatòria)
GRAU EN CIÈNCIES I TECNOLOGIES DE TELECOMUNICACIÓ (Pla 2010). (Unitat docent Optativa)
GRAU EN ENGINYERIA DE TECNOLOGIES I SERVEIS DE TELECOMUNICACIÓ (Pla 2015). (Unitat docent Optativa)
Crèdits ECTS: 6 Idiomes docència: Castellà

Professorat

Responsable: Antonio Bonafonte
Altres: Antonio Bonafonte, Asunción Moreno, José Adrián Rodríguez Fonollosa

Horari d'atenció

Horari:

Capacitats prèvies

Requisits

Competències de la titulació a les quals contribueix l'assignatura

Transversals:

07 AAT N3. APRENTATGE AUTÒNOM - Nivell 3: Aplicar els coneixements assolits a la realització d'una tasca en funció de la pertinència i la importància, decidint la manera de dur-la a terme i el temps que cal dedicar-hi i seleccionant-ne les fonts d'informació més adequades.

Metodologies docents

Classes expositives
Classes laboratori
Treball en grup (no presencial)
Treball individual (no presencial)
Altres activitats
Exercicis
Presentació oral
Proves de resposta llarga (Control)
Proves de resposta llarga (Examen Final)
Pràctica de laboratori

Objectius d'aprenentatge de l'assignatura

230022 - PAV - Processament d'Àudio i Veu

L'objectiu del curs és que l'alumne aprengui o aprofundeixi en les tècniques més habituals per al processat de veu i d'àudio així com la seva utilització en problemes d'anàlisi, classificació, síntesi o altre tipus de processat.

Resultat de l'aprenentatge:

L'assignatura contribueix al resultat de l'aprenentatge específicament en l'aspecte de processat d'informació d'àudio i veu:

Té capacitat de construir, explotar i gestionar serveis i aplicacions de telecomunicacions, enteses aquestes com sistemes de captació, tractament analògic i digital, codificació, transport, representació, processat, emmagatzematge, reproducció, gestió i presentació de serveis audiovisuals i informació multimèdia.

Està familiaritzat amb l'anàlisi, especificació, operació i manteniment de sistemes, equips, capçaleres i instal·lacions de televisió i àudio, tant en entorns fixos com mòbils.

És capaç de realitzar projectes de locals i instal·lacions destinats a la producció i enregistrament de senyals d'àudio.

Té capacitat per a crear, codificar, gestionar, difondre i distribuir continguts multimèdia, atenent a criteris de usabilitat i accessibilitat dels serveis audiovisuals, de difusió i interactius.

En els treballs en equip, planifica i acorda els objectius, les regles de funcionament, les responsabilitats, l'agenda i el procediment de revisió del treball.

Porta a terme les tasques a partir de les orientacions bàsiques donades pel professorat, decidint el temps i els recursos necessaris. Avaluja les pròpies fortaleses i debilitats i actua en conseqüència.

Hores totals de dedicació de l'estudiantat

Dedicació total: 150h	Hores grup gran:	39h	26.00%
	Hores grup petit:	26h	17.33%
	Hores aprenentatge autònom:	85h	56.67%

230022 - PAV - Processament d'Àudio i Veu

Continguts

<p>Tema 1. Introducció al processat de veu i àudio</p>	<p>Dedicació: 2h 30m Grup gran/Teoria: 2h Aprentatge autònom: 0h 30m</p>
<p>Tema 2. El senyal de veu</p>	<p>Dedicació: 6h 05m Grup gran/Teoria: 2h Aprentatge autònom: 4h 05m</p>
<p>Descripció: El llenguatge humà Producció de la parla Definició i anàlisi de característiques del parla</p>	
<p>Tema 3. Representacions de senyals de veu i àudio</p>	<p>Dedicació: 21h 20m Grup gran/Teoria: 5h 20m Aprentatge autònom: 16h</p>
<p>Descripció: Representacions basades en STFT (short-time Fourier transform) Representacions basades en LPC (predicció lineal) Processat homomorfic: el cepstrum Estimació pitch</p>	
<p>Tema 4. Robustesa respecte a l'entorn acústic</p>	<p>Dedicació: 25h 25m Grup gran/Teoria: 8h Aprentatge autònom: 17h 25m</p>
<p>Descripció: Eco i reverberació Arrays de micròfons Substracció espectral</p>	

230022 - PAV - Processament d'Àudio i Veu

<p>Tema 5. Fonaments de la classificació de veu i àudio</p>	<p>Dedicació: 16h Grup gran/Teoria: 8h Aprentatge autònom: 8h</p>
<p>Descripció: Nocions de classificadores: regressió lineal, CART (clasificacions and regression trees) Clasificadors bayesianos. GMM (gaussian mixture models) HMM (Hidden Markov Models) Detecció de veu Detecció d'esdeveniments acústics Reconeixement del parlant</p>	
<p>Tema 6. Reconeixement de la parla</p>	<p>Dedicació: 16h Grup gran/Teoria: 8h Aprentatge autònom: 8h</p>
<p>Descripció: Modelatge acústic Modelatge del llenguatge</p>	
<p>Tema 7. Síntesi de la parla</p>	<p>Dedicació: 25h Grup gran/Teoria: 12h Aprentatge autònom: 13h</p>
<p>Descripció: Processat del text Descripció i modelatge de la prosòdia Generació paramètrica. Model de formants. Generació de veu mitjançant concatenació. Mètodes PSOLA i HSM. Generació de veu mitjançant selecció d'unitats. Generació paramètrica mitjançant HMM.</p>	

230022 - PAV - Processament d'Àudio i Veu

Projecte: desenvolupament de sistema de processat de veu i/o àudio

Dedicació: 30h

Grup petit/Laboratori: 13h

Aprentatge autònom: 17h

Descripció:

En grups s'haurà de dissenyar i desenvolupar un sistema de processat de veu i àudio.

Exemple: millora d'un sistema de reconeixement de la parla, intervenint en diferents aspectes. Tots els projectes es presenten en l'aula a professors i estudiants.

230022 - PAV - Processament d'Àudio i Veu

Planificació d'activitats

Exercicis	
Descripció: Els alumnes realitzen individualment 5 treballs	
Proves de resposta llarga (Control)	
Descripció: Exàmens parcials	
Presentació oral	
Descripció: Presentació del projecte de curs	
Pràctica de laboratori	Dedicació: 0h Grup petit/Laboratori: 0h
Descripció: Projecte: desenvolupament de sistema de processat de veu i/o àudio	
Altres activitats	Dedicació: 0h Grup petit/Laboratori: 0h
Descripció: Projecte: desenvolupament de sistema de processat de veu i/o àudio	
Altres activitats	Dedicació: 0h Grup petit/Laboratori: 0h
Descripció: Projecte: desenvolupament de sistema de processat de veu i/o àudio	
Proves de resposta llarga (Examen Final)	Dedicació: 0h Grup gran/Teoria: 0h
Descripció: Examen final	

230022 - PAV - Processament d'Àudio i Veu

Sistema de qualificació

Treballs realitzats per l'alumne al laboratori i de forma autònoma (50%)
Dos exàmens parcials (50%)

En cas de no aprovar-se el curs, exàmen final.

En aquesta assignatura s'avaluaran les competències genèriques:

- Treball en equip (Nivell Mitjà)
- Aprenentatge autònom (Nivell Mitjà)

Normes de realització de les activitats

Bibliografia

Bàsica:

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.