

**FITXA IDENTIFICATIVA****DADES DE L'ASSIGNATURA****Codi:** 34907**Nom:** Processat digital d'àudio i veu**Cicle:** Grau**Crèdits ECTS:** 6**Curs acadèmic:** 2025-26**TITULACIONS**

Titulació	Centre	Curs	Període
1403 - Grau d'Enginyeria Telemàtica	Escola Tècnica Superior d'Enginyeria	4	Segon quadrimestre

MATÈRIES

Titulació	Matèria	Caràcter
1403 - Grau d'Enginyeria Telemàtica	Optativitat	OPTATIVA

COORDINACIÓ

COBOS SERRANO MAXIMO

RESUM

L'assignatura Processat Digital d'Àudio i Veu és una assignatura de quart curs que forma part de l'oferta en optativitat del Grau en Enginyeria Telemàtica. L'assignatura complementa els continguts vistos en altres assignatures del Grau com a Processament Digital del Senyal, Senyals i Sistemes Lineals i Fonaments Matemàtics de les Comunicacions, oferint una visió aplicada dels conceptes estudiats al llarg d'aquestes matèries. Així, els temes abastats per aquesta assignatura estan orientats a l'aplicació del tractament digital de senyals i l'aprenentatge automàtic en el camp del processament d'àudio i de veu.

L'assignatura justifica la importància del processament digital de senyals d'àudio en els sistemes multimèdia actuals, revisant breument alguns conceptes bàsics estudiats en cursos anteriors. L'estructuració dels continguts segueix un enfocament en el qual es comença introduint els conceptes més teòrics en la primera part del curs relacionats amb el processament digital de senyals, donant pas a altres temes més enfocats a l'aprenentatge màquina, els quals han adquirit una gran importància en l'actualitat. Així, es parteix d'una revisió de conceptes bàsics de processament de senyal, introduint diverses representacions dels senyals d'àudio, posant especial atenció a les representacions temps-freqüència. La



part teòrica es complementa amb l'estudi del sistema auditiu humà i el seu impacte en el disseny de sistemes pràctics de codificació d'àudio amb pèrdues, els fonaments del sistema fonador i els models font/filtre utilitzats en codificació de veu.

La segona part del curs introdueix els fonaments dels sistemes d'àudio basats en aprenentatge màquina, introduint l'ús de característiques utilitzades àmpliament en problemes d'àudio i la seua combinació amb classificadors binaris clàssics com les màquines de vectors suport o regressió logística. Finalment s'introdueix a l'alumne als sistemes d'aprenentatge profund actuals.

Nota: en aquesta Guia, on diu estudiant, s'entén també l'estudiant, així com alumne també alumna, professor també professora, o viceversa.

CONEIXEMENTS PREVIS

RELACIÓ AMB ALTRES ASSIGNATURES DE LA MATEIXA TITULACIÓ

No s'ha especificat restriccions de matrícula amb altres assignatures del pla d'estudis.

ALTRES TIPUS DE REQUISITS

COMPETÈNCIES / RESULTATS D' APRENTATGE

DESCRIPCIÓ DE CONTINGUTS

0. Introducció

Visió global del processament digital de senyal d'àudio, imatges i vídeo. El processament de senyal en sistemes multimèdia.

1. Revisió del processament del senyal

Introducció. Revisió de convolució i filtrat. Teoria del mostreig. Processament de senyal en temps discret. DFT i DTFT. Processos aleatoris.



2. Audio i principis psicoacústics

Descripció de contenidos (Valencià):

Introducció. Fonaments d'acústica. Audició i percepció. Sonoritat i bandes crítiques.

3. Introducció a l'aprenentatge automàtic en àudio

Processament de senyal i intel·ligència artificial en àudio. Evolució i història dels sistemes d'àudio basats en aprenentatge automàtic. Reptes relacionats amb l'ús de les tècniques d'aprenentatge automàtic en àudio. Aplicacions actuals.

4. Representacions d'àudio per l'aprenentatge automàtic

Representacions d'àudio temps-freqüència: STFT, Espectrograma Mel, Bancs de filtres Gammatone, Chromagrama. Característiques clàssiques: temporals, espectrals, MFCCs. Representacions compactes i reducció de dimensionalitat: NMF, PCA.

5. Tècniques d'aprenentatge automàtic aplicades a àudio i veu

Introducció a l'aprenentatge supervisat i no supervisat. Classificadors binaris lineals: SVM, Regressió logística. Classificadors no lineals: k-NN, models Kernel. Models de mesclades de Gaussians (GMM), k-Means. Introducció a les xarxes neuronals i l'aprenentatge profund.

VOLUM DE TREBALL (HORES)

ACTIVITATS PRESENCIALS

Activitat	Hores
Teoria	30,00
Pràctiques a l'aula	10,00
Laboratori	20,00
Total hores	60,00

ACTIVITATS NO PRESENCIALS

Activitat	Hores
Assistència a altres activitats	0,00
Elaboració de treballs individuals o en grup	30,00
Estudi i treball autònom	0,00
Preparació de classes	25,00
Preparació d'activitats d'avaluació	20,00
Resolució de casos pràctics	15,00
Total hores	90,00



METODOLOGIA DOCENT

1.- Treball presencial format per:

a) Classes de teoria, les quals consistiran en la presentació i explicació bàsica de la matèria corresponent. Es proposaran activitats de curta duració, les quals exigiran la intervenció dels alumnes amb l'objectiu de confirmar la comprensió de la teoria exposada.

b) Classes d'exercicis i seminaris pràctics, dissenyades per a resoldre problemes de major envergadura o ben temporal o ben conceptual.

c) Classes de laboratori, pensades per a comprovar experimentalment algunes de les qüestions més rellevants vistes en les classes de teoria.

2.- Treball no presencial format per:

a) Preparació de qüestionaris de tema a realitzar després de finalitzar cada unitat didàctica.

b) Preparació de les pràctiques de laboratori, per a les quals l'alumne haurà d'haver llegit i assimilat el contingut del butlletí de pràctiques, així com haver repassat la teoria rellevant.

Tutories individuals i/o col·lectives. S'estableixen unes determinades hores de tutories no programades per setmana a les quals els alumnes podran assistir per a aclarir els seus dubtes (presencialment o en manera en línia).



AVALUACIÓ

L'avaluació de l'assignatura es basarà en un sistema d'avaluació contínua que considera el treball de l'alumne al llarg del curs, complementat amb una prova final que mesurarà el coneixement global adquirit.

Els components i la seua ponderació són els següents:

- Qüestionaris per tema (realitzats després de la finalització de cada unitat didàctica): 45%
- Prova pràctica final (que avaluarà els coneixements obtinguts en les sessions de laboratori): 20%
- Qüestionari global final: 25%
- Assistència, atenció i participació en classe: 10%

La còpia o plagi manifest de qualsevol activitat que forma part de l'avaluació suposarà la impossibilitat de superar l'assignatura, sotmetent-se seguidament als procediments disciplinaris oportuns indicats en el PROTOCOL D'ACTUACIÓ DAVANT PRÀCTIQUES FRAUDULENTES EN LA UNIVERSITAT DE VALÈNCIA (ACGV 123/2020).

BIBLIOGRAFIA

- Aurélien Géron, Hands-on Machine Learning with Scikit-Learn, Keras & Tensorflow, OReilly, 2020.
- Virtanen, Plumbley, Ellis, (eds): Computational Analysis of Sound Scenes and Events, Springer, 2018.
- Giannakopoulos, T. And Pirkakis, 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
- -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