COURSE DATA

DATA SUBJECT

Code: 34907
Name: Digital speech and audio processing
Cycle: Undergraduate Studies
ECTS Credits: 6
Academic year: 2026-27

STUDY (S)

Degree	Center	Acad. year	Period
1403 - Degree in Telematics Engineering	Escola Tècnica Superior d'Enginyeria	4	Second quarter

SUBJECT-MATTER

Degree	Subject-matter	Character
1403 - Degree in Telematics Engineering	Optional subjects	ELECTIVES

COORDINATION

COBOS SERRANO MAXIMO

SUMMARY

The course "Digital Audio and Speech Processing" is a fourth year course that is part of the elective offer of the Degree in Telematics Engineering. The course complements the contents seen in other subjects of the Degree such as "Digital Signal Processing", "Linear Systems and Signals " and "Mathematical Fundamentals of Communications", offering an applied vision of the concepts studied throughout these subjects. Thus, the topics covered by this course are oriented to the application of digital signal processing and machine learning in the field of audio and speech processing.

The course justifies the importance of digital audio signal processing in current multimedia systems, briefly reviewing some basic concepts studied in previous courses. The structuring of the contents follows an approach that begins by introducing the more theoretical concepts in the first part of the course related to digital signal processing, pacing the way to other topics more focused on machine learning, which have become very important nowadays. Thus, we start with a review of basic concepts of signal processing, introducing several representations of audio signals, paying special attention to time-frequency representations. The theoretical part is complemented with the study of the human auditory system and its impact on the design of practical lossy audio coding systems, the fundamentals of the speech production system and the source/filter models used in speech coding.

The second part of the course introduces the fundamentals of machine learning based audio systems, introducing the use of features widely used in audio problems and their combination with classical binary classifiers such as support vector machines or logistic regression. Finally, the student is introduced to current deep learning systems.

PREVIOUS KNOWLEDGE

RELATIONSHIP TO OTHER SUBJECTS OF THE SAME DEGREE

There are no specified enrollment restrictions with other subjects of the curriculum.

OTHER REQUIREMENTS

COMPETENCES / LEARNING OUTCOMES

DESCRIPTION OF CONTENTS

0. Introduction

Digital audio, image and video processing overview. Signal processing in multimedia.

1. Signal processing review

Descripción de contenidos (English):
Introduction. Review of convolution and filtering. Sampling theorem. Discrete-time signal processing. DFT and DTFT. Random processess.

2. Audio and psychoacoustic principles

Descripción de contenidos (English):
Introduction. Fundamentals of Acoustics. Hearing and perception. Loudness and critical bands.

3. Introduction to machine learning in audio

Signal processing and artificial intelligence in audio. Evolution and history of machine learning based audio systems. Challenges related to the use of machine learning techniques in audio. Current applications.

4. Audio representations for machine learning

Signal processing and artificial intelligence in audio. Evolution and history of machine learning based audio systems. Challenges related to the use of machine learning techniques in audio. Current applications.

5. Machine learning techniques for audio and speech applications

Introduction to supervised and unsupervised learning. Linear binary classifiers: SVM, Logistic regression. Nonlinear classifiers: k-NN, Kernel models. Gaussian Mixture Models (GMM), k-Means. Introduction to neural networks and deep learning.

WORKLOAD

PRESENCIAL ACTIVITIES

Activity	Hours
Theory	30,00
Laboratory	20,00
Classroom practices	10,00
Total hours	**60,00**

NON PRESENCIAL ACTIVITIES

Activity	Hours
Attendance at other activities	0,00
Individual or group project	30,00
Independent study and work	0,00
Preparation of lessons	25,00
Preparation for assessment activities	20,00
Resolution of case studies	15,00
Total hours	**90,00**

TEACHING METHODOLOGY

1.- Classroom work consisting of:

a) Theory classes, which will consist of the presentation and basic explanation of the corresponding subject. Short activities will be proposed, which will require the intervention of the students with the objective of confirming the understanding of the theory exposed.

b) Exercise and seminar classes, designed to solve problems of greater temporal or conceptual scope.

c) Laboratory classes, designed to experimentally test some of the most relevant issues seen in the theory classes.

2.- Non-attendance work consisting of:

(a) Preparation of subject questionnaires to be done after finishing each didactic unit.

b) Preparation of the laboratory practices, for which the student must have read and assimilated the content of the practice bulletin, as well as having reviewed the relevant theory.

Consulting sessions: A certain number of hours are established each week, which the students can

attend in order to solve doubts

EVALUATION

The course assessment will be based on a continuous evaluation system that considers student work throughout the course, complemented by a final exam that will measure the overall knowledge acquired.

The components and their weighting are as follows:
-Topic-based questionnaires (completed after each thematic unit): 45%
-Final practical test (evaluating knowledge gained in laboratory sessions): 20%
-Final comprehensive questionnaire: 25%
-Attendance, attention, and participation in class: 10%

Any instance of copying or manifest plagiarism in any activity forming part of the evaluation will result in the inability to pass the course. Subsequently, the appropriate disciplinary procedures indicated in the PROTOCOL FOR ACTION AGAINST FRAUDULENT PRACTICES AT THE UNIVERSITY OF VALENCIA (ACGUV

123/2020) will be applied.

REFERENCES

- 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 Pikrakis, 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