

60936 - Speech technologies

Información del Plan Docente

Academic Year	2017/18
Subject	60936 - Speech technologies
Faculty / School	110 - Escuela de Ingeniería y Arquitectura
Degree	533 - Master's Degree in Telecommunications Engineering
ECTS	5.0
Year	2
Semester	First semester
Subject Type	Optional
Module	---

1.General information

1.1.Introduction

1.2.Recommendations to take this course

1.3.Context and importance of this course in the degree

1.4.Activities and key dates

2.Learning goals

2.1.Learning goals

2.2.Importance of learning goals

3.Aims of the course and competences

3.1.Aims of the course

3.2.Competences

4.Assessment (1st and 2nd call)

4.1.Assessment tasks (description of tasks, marking system and assessment criteria)

5.Methodology, learning tasks, syllabus and resources

5.1.Methodological overview

The methodology followed in this course is oriented towards achievement of the learning objectives. A wide range of teaching and learning tasks are implemented, such as lectures (P1), practice sessions (P2), supervised individual or group assignments (P3), and laboratory sessions (P4).

5.2. Learning tasks

The course includes the following learning tasks:

- **Lectures.** Presentation of the main course contents combined with the active participation of students. This activity will take place in the classroom. The theoretical contents are taught in a way that it allows students to achieve all the specified learning outcomes and competences.
- **Practice sessions.** Students solve exercises and problems in the classroom. Their contents will be closely related to the lectures. This activity is designed to gradually progress in the achievement of the learning outcomes and competences.
- **Supervised individual or group assignments.** In groups, students will solve several practical problems related to the syllabus. Solutions must be submitted in time and in the correct format. This activity is designed to consolidate all the learning outcomes and competences. The marks of these assignments is part of the assessment.
- **Laboratory Sessions** (8 hours). Sessions of 2 hours will be held in the computer classroom. For a efficient use of the sessions, some previous preparation will be required as well as some after-lab work with the obtained results to settle the concepts. Through these activities, all specified learning outcomes and competences are strengthened and reinforced. Instructions and guidelines will be provided in advance so each student will be able to find a detailed description of the activities to be performed in the lab as well as the way in which the student must show the acquisition of the relevant results and competences, since it is also an assessment task.

5.3. Syllabus

The course will address the following topics:

Topic 1. Introduction to speech technologies

- Speech Communication model
- Speech Generation
- Speech Perception

Topic 2. Fundamentals of pattern recognition

- Bayes Decision Theory
- Classifiers
- Extraction and selection of features
- Unsupervised estimation methods

Topic 3. Speech Processing

- Speech production digital model
- Short-term analysis, time-frequency representation
- Linear Prediction
- Homomorphic Analysis
- Applications: pitch and formants estimation, noise reduction

Topic 4. Automatic Speech Recognition

- Historical perspective and state of the art
- Acoustic Modelling
- Language Modelling
- Basic search algorithms
- Applications and toolkits

Topic 5. Text to speech conversion

- Historical perspective and state of the art

60936 - Speech technologies

- Linguistic Analysis
- Speech Synthesis
- Applications and toolkits

Topic 6. Biometric Speaker recognition

- Historical perspective and state of the art
- Verification and speaker identification
- Applications and toolkits

5.4.Course planning and calendar

Further information concerning the timetable, classroom, office hours, assessment dates and other details regarding this course, will be provided on the first day of class or please refer to the EINA website.

5.5.Bibliography and recommended resources

- 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
- 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