

Course guide

230623 - SLPDL - Speech and Language Processing with Deep Learning

Last modified: 25/05/2023

Unit in charge:	Barcelona School of Telecommunications Engineering	
Teaching unit:	739 - TSC - Department of Signal Theory and Communications.	
Degree:	MASTER'S DEGREE IN TELECOMMUNICATIONS ENGINEERING (Syllabus 2013). (Optional subject). MASTER'S DEGREE IN ADVANCED TELECOMMUNICATION TECHNOLOGIES (Syllabus 2019). (Optional subject).	
Academic year: 2023	ECTS Credits: 5.0	Languages: English

LECTURER

Coordinating lecturer:	Consultar aquí / See here: https://telecos.upc.edu/ca/estudis/curs-actual/professorat-responsables-coordinadors/responsables-assignatura
Others:	Consultar aquí / See here: https://telecos.upc.edu/ca/estudis/curs-actual/professorat-responsables-coordinadors/professorat-assignat-idioma

PRIOR SKILLS

Calculus and Linear Algebra
Probability and Statistics
Algorithmics and programming
Signal Processing
Large programming experience in Python, Matlab or C++.
Machine Learning
Deep Learning
Large programming experience in Python, Matlab or C++.

TEACHING METHODOLOGY

Lectures (slides) and an external invited talk (mandatory attendance)
Theoretical and practical assignments grouped into subjects

LEARNING OBJECTIVES OF THE SUBJECT

The course is focused on the study of the technologies for the development of advanced applications of spoken and written language. It focuses on new technologies based on deep learning and its application to automatic speech recognition, text to speech, and machine translation.

The final project gives students additional information about a particular topic, and also aims to help boost their own skills in the development of applications or in research.

STUDY LOAD

Type	Hours	Percentage
Hours large group	39,0	31.20
Self study	86,0	68.80

Total learning time: 125 h

CONTENTS

1. Introduction to speech processing

Description:

Introduction to speech technology and applications.
Fundamentals of speech perception and speech production.
Speech spectrum: STFT, Spectrogram.
Features: Cepstrum, MFCC, Pitch.
Techniques: Vector Quantizers, Gaussian Mixture Models.

Full-or-part-time: 6h

Theory classes: 6h

2. Deep Learning

Description:

Introduction to Deep Learning techniques and their application to speech and language processing.

Specific objectives:

Introduction to deep learning, neural nets, learning algorithms, momentum and rmsprop, regularization, word embeddings, recursive neural networks. Applications to Natural Language Processing.

Full-or-part-time: 9h

Theory classes: 9h

3. Automatic speech recognition

Description:

Pattern matching. Dynamic time warping
Hidden Markov models. Isolated word recognition
Large vocabulary continuous ASR: Acoustic modeling. Language modeling
Deep Learning for language modelling and automatic speech recognition.
Toolkits

Full-or-part-time: 10h

Theory classes: 10h



4. Speech synthesis

Description:

Linguistic processing. Prosody modeling
Waveform generation.
Concatenation and statistical methods.
Deep learning in speech synthesis.

Full-or-part-time: 7h

Theory classes: 7h

5. Machine Translation

Description:

Introduction to Machine Translation. Statistical Machine Translation
Neural Machine Translation
Speech translation.

Full-or-part-time: 7h

Theory classes: 7h

GRADING SYSTEM

Midterm exam: 20%, Final exam: 20%, Assignments: 60%

BIBLIOGRAPHY

Basic:

- Goodfellow, I.; Bengio, Y.; Courville, A. Deep Learning [on line]. Cambridge, Massachusetts: MIT Press, 2016 [Consultation: 22/06/2016]. Available on: <http://www.deeplearningbook.org/>. ISBN 9780262035613.
- 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.