

## 230623 - SLPDL - Speech and Language Processing with Deep Learning

Coordinating unit:	230 - ETSETB - Barcelona School of Telecommunications Engineering
Teaching unit:	739 - TSC - Department of Signal Theory and Communications
Academic year:	2019
Degree:	MASTER'S DEGREE IN ADVANCED TELECOMMUNICATION TECHNOLOGIES (Syllabus 2019). (Teaching unit Optional) MASTER'S DEGREE IN TELECOMMUNICATIONS ENGINEERING (Syllabus 2013). (Teaching unit Optional) MASTER'S DEGREE IN INFORMATION AND COMMUNICATION TECHNOLOGIES (Syllabus 2009). (Teaching unit Optional)
ECTS credits:	5
Teaching languages:	English

### Teaching staff

Coordinator:	Rodriguez Fonollosa, Jose Adrian
Others:	Rodriguez Fonollosa, Jose Adrian Bonafonte Cavez, Antonio Jesus Ruiz Costa-Jussa, Marta

### Prior skills

Calculus and Linear Algebra  
Probability and Statistics  
Algorithmics and programming  
Signal Processing

### Requirements

Large programming experience in Python, Matlab or C++.

### Teaching methodology

Lectures (slides) and an external invited talk (mandatory attendance)  
Theoretical and practical assignments grouped in subjects  
Research project, presented in written and oral form

### 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.



## 230623 - SLPDL - Speech and Language Processing with Deep Learning

### Study load

Total learning time: 125h	Hours large group:	39h	31.20%
	Hours medium group:	0h	0.00%
	Hours small group:	0h	0.00%
	Guided activities:	0h	0.00%
	Self study:	86h	68.80%

## 230623 - SLPDL - Speech and Language Processing with Deep Learning

### Content

1. Introduction to speech processing	Learning time: 6h Theory classes: 6h
<p>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.</p>	
2. Deep Learning	Learning time: 9h Theory classes: 9h
<p>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.</p>	
3. Automatic speech recognition	Learning time: 10h Theory classes: 10h
<p>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</p>	
4. Speech synthesis	Learning time: 7h Theory classes: 7h
<p>Description: Linguistic processing. Prosody modeling Waveform generation. Concatenation and statistical methods. Deep learning in speech synthesis.</p>	

## 230623 - SLPDL - Speech and Language Processing with Deep Learning

5. Machine Translation	Learning time: 7h Theory classes: 7h
Description: Introduction to Machine Translation. Statistical Machine Translation Neural Machine Translation Speech translation.	

### Qualification system

Midterm exam: 25%, assignments: 30%, research work: 45%

### Bibliography

Basic:

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.

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.