Course guides
230362 - DLSL - Deep Learning for Speech and Language

Last modified: 06/05/2019

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).

Academic year: 2019 **ECTS Credits:** 2.5 **Languages:** English

LECTURER

Coordinating lecturer: Giró Nieto, Xavier

Others: Bonafonte Cavez, Antonio Jesus
Rodriguez Fonollosa, Jose Adrian
Ruiz Costa-Jussa, Marta
Hernando Pericas, Francisco Javier
Pascual, Santiago
Sayrol Clols, Elisa
Giró Nieto, Xavier

PRIOR SKILLS

It is advisable to have some background in machine learning and signal processing. Students will also develop their projects in Python, so previous contact with this language is recommended.

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
CE1. Ability to apply information theory methods, adaptive modulation and channel coding, as well as advanced techniques of digital signal processing to communication and audiovisual systems.

Transversal:
CT3. TEAMWORK: Being able to work in an interdisciplinary team, whether as a member or as a leader, with the aim of contributing to projects pragmatically and responsibly and making commitments in view of the resources that are available.

CT4. EFFECTIVE USE OF INFORMATION RESOURCES: Managing the acquisition, structuring, analysis and display of data and information in the chosen area of specialisation and critically assessing the results obtained.

CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

TEACHING METHODOLOGY

- Lectures
- Application classes
- Group work
- Group work (distance)

LEARNING OBJECTIVES OF THE SUBJECT

The aim of this course is to train students in methods of deep learning for speech and language. Recurrent Neural Networks (RNN) will be presented and analyzed in detail to understand the potential of these state of the art tools for time series processing. Engineering tips and scalability issues will be addressed to solve tasks such as machine translation, speech recognition, speech synthesis or question answering. Hands-on sessions will provide development skills so that attendees can become competent in contemporary data anlaytics tools.

STUDY LOAD

Type	Hours	Percentage
Hours small group	10,0	16.00
Self study	42,5	68.00
Hours large group	10,0	16.00

Total learning time: 62.5 h

CONTENTS

Advanced Neural Networks

Description:
- Architectures: LSTM, GRU, recursive, Conv-LSTM, dynamic memory networks, TDNN, highway networks...
- Training: datasets, back-propagation, optimization, adversarial...
- Learning: supervised/unsupervised, continual...
- Visualization
- Attention models
- Embeddings: seq2seq, skip-though vectors...
- Ensembles of models

Specific objectives:
At the end of this course, students will be familiar with the state of the art techniques based on deep learning architectures.

Full-or-part-time: 16 h
Theory classes: 6h
Guided activities: 8h
Self study : 2h 30m

Applications to Speech and Language

Description:
- Natural Language Processing
- Machine Translation
- Speech recognition
- Speaker recognition
- Speech synthesis
- Multimodal: language and vision.
- Frameworks and tools: TensorFlow, Keras, Kaldi

Full-or-part-time: 50 h
Theory classes: 6h
Guided activities: 10h
Self study : 34h

GRADING SYSTEM

Lectures: 30%
Practical: 60%
Attendance: 10%

BIBLIOGRAPHY

Basic:
- Socher, Richard. Deep Learning for Natural Language Processing [on line]. Stanford University, 2016 [Consultation: 05/09/2016]. Available on: http://cs224d.stanford.edu/.
- Goodfellow, I.; Bengio, Y.; Courville, A. Deep Learning [on line]. MIT Press, 2017 [Consultation: 05/09/2016]. Available on: http://www.deeplearningbook.org/. ISBN 0262035618.

Complementary:
- Joan Bruna. Topics Course on Deep Learning [on line]. UC Berkeley, 2016 [Consultation: 05/09/2016]. Available on: https://github.com/joanbruna/stat212b.
- Dhruv Batra. Deep Learning for Perception [on line]. Blacksburg, VA, USA: Virginia Tech, 2016 [Consultation: 02/05/2020]. Available on: https://computing.ece.vt.edu/~f15ece6504/.
- Aaron Courville and Yoshua Bengio. Deep Learning Summer School [on line]. Montreal, Quebec: CIFAR, ICRA, CRM, 2016 [Consultation: 05/09/2016]. Available on: https://sites.google.com/site/deeplearningsummerschool2016/home.
- Luong, T.; Cho, T.; Manning, C. Neural Machine Translation [on line]. ACL Tutorials, 2016 [Consultation: 05/09/2016]. Available on: https://sites.google.com/site/acl16nmt/.
- Giró-i-Nieto, X.; Sayrol, E.; Salvador, A.; Torres, J.; Mohedano, E.; McGuinness, K. Deep Learning for Computer Vision [on line]. Barcelona: UPC, 2016 [Consultation: 05/09/2016]. Available on: http://imatge-upc.github.io/telecombcn-2016-dlcv/.

RESOURCES

Other resources:
The details of this course are available and updated online at: https://telecombcn-dl.github.io/2017-dlsl/