Last update: 06-05-2019

230374 - NLPDL - Natural 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 TELECOMMUNICATIONS ENGINEERING (Syllabus 2013). (Teaching unit Optional)
MASTER'S DEGREE IN ADVANCED TELECOMMUNICATION TECHNOLOGIES (Syllabus 2019). (Teaching unit Optional)

ECTS credits: 3 Teaching languages: English

Teaching staff

Coordinator: Marta Ruiz Costa-jussà

Others: José Adrián Rodríguez Fonollosa
Noé Casas
Xavi Giró

Opening hours

Timetable: During the course, after classes, from 14h to 15h

Prior skills

Calculus and Linear Algebra
Probability and Statistics
Algorithmics and programming
Natural language processing

Requirements

Student will benefit from experience in programming in Python

Degree competences to which the subject contributes

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
-Programming
-Invited talk
-Research project

Last update: 06-05-2019

230374 - NLPDL - Natural Language Processing with Deep Learning

Learning objectives of the subject

The course is focused on the study of how Deep learning techniques are applied to Natural Language Processing (NLP). The introduction of the course includes an overview of classical neural networks structures such as convolutional neural networks and recurrent neural networks and it motivates their current success in NLP. The course also reviews the core tasks of NLP (like preprocessing, part-of-speech tagging and parsing) and how they are successfully addressed with neural networks architectures. Then, the course provides a general overview of relevant NLP tasks (e.g. language modeling, neural machine translation) and it gives special focus on the following NLP applications: text classification, natural language inference, text summarization, dialog systems, which have a great impact in nowadays society both at the industry and academic level.

Finally, the course focuses on recent transversal topics. Some of these transversal topics include bias in NLP, referring to the fact that current NLP systems suffer from gender or race biases. A second transversal topic are the current latest and powerful neural architectures that are starting to be applied successfully in NLP such as adversarial networks that put two neural networks to compete to better perform in a task. Finally, the third topic from this section includes multimodal tasks such as visual question answering that use image and text at the same time.

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

Total learning time: 75h	Hours large group:	16h	21.33%
	Hours small group:	8h	10.67%
	Self study:	51h	68.00%

Last update: 06-05-2019

230374 - NLPDL - Natural Language Processing with Deep Learning

Content

Natural Language Processing with Deep Learning	Learning time: 24h Theory classes: 16h Laboratory classes: 8h

Description:

1. Introduction to Natural Language Processing (2h lecture)
This lesson includes:
o NLP definition, text processing & tasks
o Approaches (rule-based, statistical, ML, with Deep learning)
o Linguistics levels: morphology, syntax, semantics & pragmatics

2. Basic theory: Distributional models and Deep Learning Architectures (2h lecture, 1h programming)
This lesson includes theory and practice of:
o Distributional models. Current NLP techniques are based on learning information from context, so for example, two words are more similar if they appear in similar contexts. We will describe how word embeddings based on distributional learning are trained.
o Latest Deep learning architectures for NLP: classification, sequence and sequence-to-sequence modelling (encoder-decoder, attention, rnn, cnn and transformer). Here, we will describe current neural networks architectures that are applied to NLP and put them in context of NLP tasks.

3. Core NLP tasks with Deep learning (2h lecture, 1h programming)
This lesson includes theory and practice of:
o Text preprocessing (meaning tokenization of text into words, lowercasing/truecasing and normalization). Preprocessing is necessary to clean and reduce vocabulary of huge datasets.
o Pos-Tagging is the task of identification of words as nouns, verbs, adjectives...
o Parsing is the task of decomposing and define syntactic and semantic structures in a sentence or text.
o Coreference is the task of detecting in a sentence or a text several forms/expressions that refer to the same person or thing.

4. NLP applications (5h theory, 5h programming)
These lessons are the main part of this subject and include theory and practice of relevant applications in NLP, including:
o Overview of NLP applications
o Text classification (easy task: spam/not spam, complex task: sentiment analysis)
o Natural language inference which is the task of evaluating the similarity or dissimilarity in meaning of two sentences
o Text summarization which is the task of extracting the main ideas of a document.
o Dialog Systems which is the task of automatically have a conversation with a machine.

5. Transversal NLP topics (6h theory)
These lessons includes the following topics that apply to NLP in general:
o Bias, detecting and solving bias problems in gender, race and others. Adjectives like “eminence” tend to be more associated to men and “sexy” to women. This has become a huge concern in NLP and AI applications.
o Adversarial architectures which put two neural architectures to compete
o Multimodality, combining text and image, such as visual question answering.

230374 - NLPDL - Natural Language Processing with Deep Learning

Qualification system

Communication: 20%
Assignments: 40%
Research work: 40%

Bibliography

Basic:

Goodfellow, I.; Bengio, Y.; Courville, A. Deep learning [on line]. Cambridge, Massachusetts: The MIT Press, 2016 [Consultation: 17/10/2018]. Available on: <http://www.deeplearningbook.org/>. ISBN 9780262035613.

Goldberg, Y. Neural network methods in natural language processing. Morgan & Clypool Publishers, 2017. ISBN 9781627052986.