230362 - DLSL - Deep Learning for Speech and Language

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
230362 - DLSL - Deep Learning for Speech and Language

translation, speech recognition, speech synthesis or question answering. Hands-on sessions will provide development skills so that attendees can become competent in contemporary data analytics tools.

### Study load

<table>
<thead>
<tr>
<th>Study load</th>
<th>Hours large group:</th>
<th>Hours small group:</th>
<th>Self study:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total learning time:</strong> 62h 30m</td>
<td>10h</td>
<td>10h</td>
<td>42h 30m</td>
</tr>
<tr>
<td><strong>Learning time:</strong> 16h 30m</td>
<td>16.00%</td>
<td>16.00%</td>
<td>68.00%</td>
</tr>
</tbody>
</table>

### Content

#### 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-thought 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.

#### 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

**Learning time:** 50h
- Theory classes: 6h
- Guided activities: 10h
- Self study: 34h
Qualification system

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

Bibliography

Basic:


Complementary:


Others resources:

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