|  |  |  |
| --- | --- | --- |
| **DS&AI course building blocks** | COMPUTATIONAL LINGUISTICS (3 Credits) | |
| **Section 1** | **Introduction** | |
| Lesson 1 | Introduction: challenges of text data | Slides |
| Lesson 2 | Tokenization and sentence splitting, morphology | Slides, tutorial |
| Lesson 3 | Regular expressions and edit distance | Slides |
| **Section 2** | **Language processing** | |
| Lesson 4a | POS tagging | Slides |
| Lesson 4b | Sequence labelling and recurrent neural networks | Slides, tutorial |
| Lesson 5 | Sequence-to-sequence transformation | Slides, tutorial |
| **Section 3** | **Word representations (vector semantics and embeddings)** | |
| Lesson 6 | Vector space model and classification | Slides, tutorial |
| Lesson 7 | Word embeddings | Slides, tutorial |
| **Section 4** | **Syntactic processing** | |
| Lesson 8 | Constituent grammar and parsing | Slides |
| Lesson 9 | Dependency parsing | Slides |
| **Section 5** | **Semantic analysis** | |
| Lesson 10 | Word sense disambiguation | Slides |
| Lesson 11 | Semantic role labelling | Slides |
| **Section 6** | **Classification Models** | |
| Lesson 12 | Naïve Bayes | Slides, assignment |
| Lesson 13 | Feedforward neural networks | Slides, tutorial |
| **Section 7** | **Neural NLP and transfer learning** | |
| Lesson 14 | Contextual word representations (BERT language models) | Slides, tutorial |
| **Section 8** | **Applications of Computational Linguistics** | |
| Lesson 15 | Information Extraction | Slides,assignment |
| Lesson 16 | Sentiment Analysis | Slides |
| Lesson 17 | Dialogue systems /Conversational agents | Slides, assignment |
| Lesson 18 | Machine translation | Slides |

Total 26 hours lab sessions and 19 hours lessons