

Syllabus

Week	Lecture	Lab
1	Introduction	Linux, Python, Jupyter notebooks
2	Encoding text. Regular expressions. Basic text processing	Codecs, regular expressions, spaCy, stanfordnlp
3	Morphology and Finite State Transducers	Finite State Transducers with Pynini
4	N-gram language modeling	Finite State Transducers with Pynini, cont.
5	Text classification with Naive Bayes	Text Classification with sklearn
6	Classifying words using Logistic Regression	Named Entity Recognition using sklearn and CRFSuite
7	Neural networks	Computation Graphs and Neural Networks with Pytorch
8	Neural language models and word Embeddings	Using Pre-trained Word Embeddings with Gensim and Pytorch
9	Convolutional and recurrent models for text processing	Convolutional Neural Networks for Text Classification. Recurrent Neural Networks for Name Classification
10	Neural attention, transformer models	Generating Text using Recurrent Neural Network
11	Contextual word embeddings, pretrained language models	Named entity recognition using BERT
12	Machine translation (guest lecture)	TBA
13	Speech features, end-to-end speech recognition	Speech activity detection using convolutional and recurrent models
14	Speaker and language recognition	Default independent project guide. Tips and Tricks.
15	Self-supervised pretraining for speech processing	Using wav2vec for quick finetuning of speech models