

Concepts and Applications in NLP

Organization and Overview

Marion Di Marco

October 15, 2024

General Information

- Lectures
 - introduce and discuss a new topic each week
 - look at linguistic background and NLP techniques
 - no lecture on November 12!
- Course website: <https://mariondimarco.github.io/teaching/>
- 4 homework assignments through the semester
 - objective: get more insight into some of the presented applications
 - presentation and discussion of a paper
- Material and reading: most slides are based on content from *Speech and Language Processing (3rd ed. draft)*, Dan Jurafsky and James H. Martin (2024)
<https://web.stanford.edu/~jurafsky/slp3/>

Topics

- Introduction: Linguistic Concepts
- Morphology
- Sequence Labeling and POS Tagging
- Grammars, Parsing, Structure
- Text Classification and Sentiment Analysis
- Information Extraction
- Distributional Semantics
- Machine Translation
- Semantic Role Labeling
- Language Modeling: n-gram models
- Large Language Models

Introduction: Natural Language Processing

- Human language has many layers and challenges
- Humans can understand and produce language effortlessly
 - understanding of linguistic structure
 - world knowledge
 - ability to generalize and interpret context
- What about computers?
 - good at detecting statistical patterns
- How well can computers capture language?
How well can computers handle the challenges of language?
- NLP: Combines the fields of computer science, AI and linguistics