

Introduction

(SNLP tutorial)

Vilém Zouhar, Awantee Deshpande, Julius Steuer

April 16, 2021

Overview

- Hello
- Topics (15 minutes)
- Requirements (5 minutes)
- Materials
- Assignments
- Homework (10 minutes)

Hello

Who am I?

Hello

Who am I?

Who are you?



Topics

Task: Pick one not yet taken + why do you find it interesting.

- Language properties, Zipf's Law, basic statistical formalism
- Entropy, basic information theory (Shannon's game, entropy-based quantities, code lengths)
- Language modelling, back-off models (interpolation, discounting)
- Text classification, basic algorithms (kNN, decision trees, SVM, ...)
- Word sense disambiguation, basic algorithms (dictionary-, translation-, collocation-based)
- Information retrieval, latent semantic analysis, singular value decomposition
- Machine translation, word alignment, beamsearch
- POS tagging, named entity recognition
 - ▶ sequence labeling (hidden markov chains / models, conditional random fields)

Requirements

Tutorial Requirements (exam admission)

- 70% of mandatory points (~10 assignments)
- Presenting a solution to an exercise *at least* once
- Tutorial points only for exam admission (no final grade influence)

Tutorial Bonus Points

- ~2pts for extra exercises in the assignments
- 1pt for participating and *talking* in a tutorial

Final Project

- 25% of the final grade
- Details TBD

What's available

- Lectures by prof. Klakow
- Live tutorials
- Tutorial handouts
- Consultations
 - ▶ Only in specific cases
 - ▶ **no** email and **no** chat by default
- Public forum (please use Piazza)
 - ▶ Other students will also benefit from the answers
 - ▶ You can answer someone else's issue
- Corrected homework

Assignments

- Mandatory groups of 2
- Usually 3 exercises per one assignment
 - ▶ Can't be changed later (very special exceptions)
- Jupyter notebook templates
 - ▶ Assignment + solution in the same notebook
 - ▶ Can use Google Collab or local runtime
- Only one submission per group
 - ▶ Submit through Teams

Dates / Times

- Lecture: Fridays 8:30-10:00
- Tutorials:
 - ▶ Awantee: TODO
 - ▶ Julius: TODO
 - ▶ Vilém: TODO
- Exam: (TBD) 30. Jul.

Current Homework

TBD

Resources

- ① UdS SNLP Class: <https://teaching.lsv.uni-saarland.de/snlp/>
- ② Tutorial repository for these slides: <https://github.com/zouharvi/uds-snlp-tutorial>
- ③ Piazza: <https://piazza.com/uni-saarland.de/spring2021/snlp>