Introduction

(SNLP tutorial)

Vilém Zouhar, Awantee Deshpande, Julius Steuer

April 20, 2021

Overview

- Hello
- Topics (15 minutes)
- Requirements
- Materials
- Assignments
- Homework (30 minutes)

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, 10 points each)
- Tutorial points only for exam admission (no final grade influence)

Tutorial Bonus Points

- ~2pts for extra excercises in the assignments
- 1pt for participating and *talking* in an tutorial
- Presenting a solution to an excercise (~5 points)
  - Presentable excercises are marked in the assignment sheet
  - Let individual tutors known if you wish to present (first come - first serve)
  - Every group can present *at most* once, about 10 to 15 minutes

Final Project

- 25% of the final grade
- Details TBD

Transfer from last year

- Possible
- Do project and exam

What's available

- Lectures by prof. Klakow (recorded)
- Tutorials
- Corrected homework
- Consultations
  - Only in specific cases
  - By default **no** email and **no** chat
  - Better ask during the lecture / tutorials
- Public forum (please use Piazza)
  - Ask questions
  - Other students will also benefit from the answers
  - You can answer someone else's issue

Assignments

- Mandatory groups of 2
- Usually 3 excerises 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
- Assignments
  - Release (usually) Friday 23:59
  - Deadline (next) Friday 23:59 (also in Teams)
- Exam: (TBD) 30. Jul.

Tutorial Content

- Review of the topic (per demand)
- Presentation of the past assignment
- Troubleshooting current assignment

Current Homework

- Notebook instructions
- Stick breaking
- Zipf's law on words
- Bonus: Zipf's law on characters

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