Linguistische Informatik

Sommersemester 2023

Dr. Thomas Eckart

2023-04-03

AGENDA

- Goals of Lecture
- Teaching Syllabus and Methods
- Materials / Sessions
- Exam
- Contact to lecturer team

GOALS OF LECTURE

COMPUTER SCIENCE AS A TOOL OR THEORETICAL BASIS.

- **Linguistics with the computer (Computational Linguistics)**: supporting linguists in their practical work
- **Natural Language Processing**: application of concepts and procedures of computer science for the analysis of language (as a system or single language).
- **Main Goal**: We would like to provide you with a basic theoretical and practical knowledge of the concepts that will enable you to independently apply the lecture contents.

  - Create a basic understanding

  - Ability to understand and apply natural language processing in various applications

SYLLABUS IN A NUTSHEL

LITERATURE

- D. Jurafsky, J. H. Martin. *Speech and Language Processing* Prentice Hall 2008. (see also 3rd edition)

- C. D. Manning, H. Schütze. *Foundations of Statistical Natural Language Processing* MIT Press 1999.

- J. Eisenstein *Introduction to Natural Language Processing* MIT Press 2019.

- C. Biemann, G. Heyer, U. Quasthoff. *Wissensrohstoff Text - Eine Einführung in das Text Mining*. SpringerLink 2022.
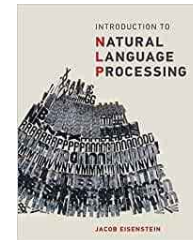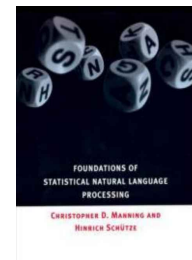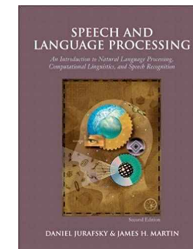
NLP CONFERENCES

- AACL - Conference of the Asian Chapter of the ACL.
- ACL - Annual Meeting of the Association for Computational Linguistics.
- COLING (linguistic focus) - International Conference on Computational Linguistics.
- CoNLL - International Conference on Natural Language Learning.
- EACL - Conference of the European Chapter of the ACL.
- EMNLP (empirical focus) - Conference on Empirical Methods in Natural Language Processing.
- NAACL - Conference of the North-American Chapter of the ACL

SOFTWARE

Industry

- SpaCy (Industrial Strength NLP), spacy.io

Research

- NLTK Natural Language Toolkit, nltk.org
- OpenNLP, opennlp.org
- Stanford NLP Group (Java, Python), nlp.stanford.edu/software
- Quanteda, quanteda.io

Machine Learning

- Gernsim, radimrehurek.com/gensim
- Scikit, scikit-learn.org
- Liblinear, csie.ntu.edu.tw/~cjlin/liblinear/
- Classification And Regression Training (Caret), topepo.github.io/caret/
- kernLab, cran.r-project.org/web/packages/kernlab/index.html
- E1071, cran.r-project.org/web/packages/e1071/index.html
- RandomForest, cran.r-project.org/web/packages/randomForest/

SYSTEM OF THE LECTURE

Course uses materials from Natural Language Processing, Machine Learning and Linguistics provided by the **TEMIR Group** and the **Computational Humanities group**. Credits are on the slides.

- A large part consists of the **theoretical lecture parts**.
- You will work with **tutorials in Python** in the exercises to learn the practical application of the techniques.
- Additionally, we will have **some theoretical exercises** in order to give deeper understanding...

MATERIALS / SESSIONS

- The materials consist of lecture slides and tutorials. You will find these materials on Moodle. General information also on TEMIR page.
- Lecture **and** exercise: **HS 12**
- Exercise participation is a prerequisite to complete the module.

- Lectures: every Monday, 13:15 - 14:45 (HS 12)
- Excercise: every Monday, 15:15 - 16:45, **starting 17.04.2023** (HS 12)

EXERCISE

TECHNICAL REQUIREMENTS

- **Python**

    – Possible IDE: VS Code, PyCharm

- **Location**:

    – HS 12 - **BYOD** (Bring your own device)

- Exercise will be held using documented and commented **Jupyter Notebooks** by the instructor

    – You can follow the code or run the programs yourself in order to understand implementation details of the theoretical concepts

EXAM

- Written exam

- Aim: during the two weeks after lecture period. Ideally at the time of the lecture.

  - It remains to be seen whether we will get a room for this time.

- Time: 60 min

  - Theoretical and exercise related questions

CONTACT

Lecture:

Thomas Eckart

Office: (SAW Leipzig)

E-Mail:

teckart@informatik.uni-leipzig.de /
  eckart@saw-leipzig.de

Exercise:

Christian Kahmann

Office: P818

E-Mail:
kahmann@informatik.uni-leipzig.de

For questions: Please make appointments with the lecturers and do not show up spontaneously at the office.