

# Seminar: Deep NLP

WS 2017 / 2018

Session 1, Wednesday October 18<sup>th</sup>, 2017  
(Introduction + Organization)

Prof. Dr. Hannah Bast  
Chair of Algorithms and Data Structures  
Department of Computer Science  
University of Freiburg

# Overview of this session

---

## ■ Demos

- Six demos involving NLP and Deep Learning

NLP = Natural Language Processing

## ■ Organizational

- Schedule, Presentations, your job, our job

## ■ Topics

- A very short introduction to each of the available topics and how it is related to the demos

On our Wiki page, you also find a list of the topics + a short description + preliminary materials

## ■ Schedule

- First week of semester: introductions for all seminars  
You can go to as many of these as you like
  - Until beg. of second week: make your choice(s) in HISinOne  
If you are interested in several seminars, give priorities
  - End of second week: assignments made in HISinOne
  - Third week: now the real seminar starts
- November 8: ML basics, part 1 + topic assignments
- November 15: ML basics, part 2
- November 22: first talks from one of the participants

...

## ■ Presentations

- For each topic, we can assign one or two persons
- Topics with one person:
  - Presentation: 30 minutes
  - Discussion: 30 minutes
- Topics with two persons:
  - Presentation 1: 20 minutes
  - Presentation 2: 20 minutes
  - Discussion: 30 minutes
- All talks should be in English + slides should be in PPT or PDF
- Each talk will be recorded (unless you object)

- Schedule for each individual presentation

Research the given topic (starting from the pointers given on the website) and make a plan of what you want to talk about

**Meeting** with us, 3 weeks before your presentation: show us your plan + we settle on the scope of your presentation

Understand / work out all the necessary details and play around (extensively) with existing software or write your own

**Meeting** with us, 2 weeks before your presentation: show us what you have done + we try to help with remaining problems

Prepare your presentation and the demos you want to show

**Meeting** with us, 1 week before your presentation: show us what you have prepared + we help with remaining problems

Finish your presentation and demo, including all the details

## ■ Your job

- You must come prepared to each of the three meetings before the presentation

If you come unprepared, the meeting will be very short

- Your presentation should be understandable and interesting

Think about all the bad presentations you had the misfortune to sit through ... you don't want to do that to other people

- You should really understand what you are talking about

Do not just repeat or paraphrase things you have read

- Your presentation should stick to the time limits

## ■ Our job

- We prepared an interesting set of topics for you
- We will meet three times with each of you
- We will help with any problems that you might have

Under the condition that you have done your share of the work ... we will not do your part

- Who are we: (see the Wiki page for contact information)

Hannah Bast (introduction today)

Niklas Schnelle (ML basics, part 1)

Thomas Goette (ML basics, part 2)

The pre-presentation meetings are with Niklas+Thomas

# Demos involving Deep NLP 1/6

## ■ Aqqu [demo](#)

- Data: a knowledge base containing information triples

Python (Lang)	language designer	Guido van Rossum
Guido van Rossum	country of nationality	Netherlands
Python (Animal)	lifespan in years	30

...

- Input: a question in (free form) natural language

who invented python

- Output: a SPARQL query on the knowledge base

```
SELECT ?x WHERE {  
  "Python (Lang)" "language designer" ?x }
```



# Demos involving Deep NLP 2/6

---

## ■ Squad [demo](#)

- Data: a piece of text

Nikola Tesla (Serbian Cyrillic: Никола Тесла; 10 July 1856 – 7 January 1943) was a Serbian American inventor, electrical engineer, mechanical engineer, physicist, and futurist best known for his contributions to the design of the modern alternating current (AC) electricity supply system.

- Input: a question in (free form) natural language

What does AC stand for?

- Output: the answer, based on what's in the text  
alternating current

# Demos involving Deep NLP 3/6

## ■ Visual QA [demo](#)

- Given a picture and a natural language question



What animals are in the picture?

How many birds?

Can I wear a T-shirt?

Is the sky visible?

- Give the (hopefully correct) answer

What animals are in the picture?

Ants

How many birds?

42

Can I wear a T-shirt?

Why not

Is the sky visible?

No

- Question Completion [demo](#)
  - Input: the beginning of some question  
who invented the in
  - Suggest continuations / completions of the question + simultaneously recognize named entities  
who invented [the internet](#)  
who invented the [integrated circuit](#)

## ■ Sentiment Analysis [demo](#)

- Given a text (think of a review on Amazon), give a score on the spectrum from negative to positive

Awesome book if your passion is photography.

I thought this book would be great, and I saw that it had rave reviews so I got it. I must say I was fairly disappointed.

- On a finer level, a review can have different shades of sentiment for different parts ... see demo

- Text repair [demo](#)

- Given a text with spelling mistakes

Tihs is a short tset sentnece with quit a feew tipos

- Compute the correct text

This is a short test sentence with quite a few typos

# List of Topics 1/3

---

- Language models (with recursive neural networks)
  - Given a sequence of words or characters, predict the next word or character
  - Used in: QA Completion, Sentiment Analysis, Text Repair
- Word vectors (e.g. Google's word2vec)
  - Representing words as vectors in a high-dimensional space
  - Used in: Question Answering, QA Completion
- Paraphrasing and synonyms
  - Identify if different words / phrases have the same meaning
  - Used in: Question Answering, QA Completion

- Convolutional Neural Networks
  - A NN structure from image processing applied to NLP
  - Used in: QA Completion, Squad, Visual QA
- Text classification
  - Classify given documents / text records into categories
  - A fundamental NLP problem
- Attention
  - Decide on which part of the (large) input to focus on
  - Used in: Visual QA

# List of Topics 3/3

---

- Question Answering on Text
  - This is what we have seen in the SQUAD demo
- Question Answering on Knowledge Bases
  - This is what we have seen in the AQQU demo
- Visual Question Answering (backup topic)
  - This is what we have seen in the VQA demo

You also find these topics + short descriptions on our Wiki page, along with some links to a preliminary selection of materials

<http://ad-wiki.informatik.uni-freiburg.de/teaching/DeepNlpWS1718>