

Embeddings for NLP and IR

Cristina España-Bonet

UdS & DFKI

Summer Semester 2019

17th April 2019

Embeddings for NLP and IR

General Info

- Questions on registration, credits and so on:
 - CS: studium@cs.uni-saarland.de
 - CoLi: Dr.Thater / Ms.Deeg
- Credits
 - CS: only seminars with 7 CP's exist
(so *must* do the term paper)
 - CoLi: can chose 4 and 7 CP options
- The registration in HISPOS is the same for all students with deadline July 3rd.

Embeddings for NLP and IR

Problem: 15 students limit

Before attending, answer these questions

- Are there more seminars that also interest you?
- Are you in your 1st semester?
Next year you'll take more benefit of the course
- Do you know what a FFN is?
The course will be more technical
- Wanna work more than you were told last week?
Evaluation through questionnaires too

Embeddings for NLP and IR

<https://cristinae.github.io/teaching/we4nlp19/>

Structure (12 Weeks)

- **Block I:** Standard introductory lectures (*me!*)
- **Block II:** Basic/SotA Models (*you!*)
word2vec, glove, fasttext, ELMo, OpenAI-GTP, BERT
- **Block III:** Advanced Models/Applications (*you!*)

Evaluation

- 70% talks (2 per student) +
30% interaction/**questionnaires**
- Term paper (topic to be agreed)

Embeddings for NLP and IR

Block I

Lectures (Intro + 3 Weeks)

- **L1:** Introduction to Word Embeddings
- **L2:** Demo/short exercise
Bring your laptop with python, anaconda, jupyter notebook, gensim, matplotlib
- **L3:** Introduction to PCA, SVD an others
- **L3:** How to do a scientific talk

Later in the course

- **LX:** How to write a scientific/term (paper) and term paper topic discussion

Embeddings for NLP and IR

Block II

Main Topics (2h*3 Weeks)

- **L5:** Word2Vec
- **L6:** Glove & FastText
- **L7:** ELMo, OpenAI-GTP, BERT

Distribution

- 5 people per group
- A group presents a topic in 2h
- Including: given papers, related papers, demos, blogs, etc..

Embeddings for NLP and IR





Block II

Let's try the 15 students limit... with a doodle.

Presentation for Block II

by Cristina • seconds ago • Print

 Building C7.2, Room -1.05

| | | | |
|--------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|
| | Word2Vec -- 29th May -- 8:30-10:30 | Glove & FastText -- 5 June -- 8:30-10:30 | Bert & Co. -- 12/19 June (TBC) -- 8:30-10:30 |
| 0 participants | ✓ 0 | ✓ 0 | ✓ 0 |
|  <input type="text" value="Enter your name"/> |  |  |  |

Embeddings for NLP and IR

Block III

- We'll discuss later