

Call for Master Thesis

Embedding of Temporal Information for Named Entity Disambiguation

Are you interested in making a big impact with your thesis? Work with us on an innovative approach for NED by embedding temporal information.

A named entity mention in natural language text, as e.g. “the president”, may refer to multiple entities, and the process of resolving the appropriate meaning in context is called **Named Entity Disambiguation** (NED). It is the task of linking a named entity mention in the textual document to an instance in a knowledge base, typically Wikipedia-derived resources like DBpedia[1]. On the other hand, embedding provides a low dimensional space in which you can translate high dimensional vectors. For instance, word embedding (a type of embedding) is a learned representation of text where the words with same or similar meaning have similar representation leveraging the semantic similarity of the words in the text.

In this thesis, your focus will be on the disambiguation of the named entity mention in the text based on the embedding of the temporal information. Temporal embedding refers to the creation of the vector representation of the temporal information present in the text. The intuition behind this work is that named entity mentions in the text which share a common time frame should have a similar vector representation and will appear closer to each other in the vector space.

The aim of this thesis is to develop a NED approach using temporal embedding. The students will use DBpedia as well as Wikipedia articles to extract the temporal information for each entity. This temporal information together with the entities is to be put together in form of a network followed by an embedding approach. Possible approaches will have to use different types of embedding approaches (e.g. Word2Vec, node2vec, RDF2Vec etc.) to generate the embeddings followed by Machine Learning based approaches to perform the Named Entity Disambiguation.

This thesis will be supervised by **Prof. Dr. Harald Sack, Information Service Engineering at Institute AIFB, KIT, in collaboration with FIZ Karlsruhe.**

[1] <http://wiki.dbpedia.org/>



Which prerequisites should you have?

- Good programming skills in Python or Java
- Interest in Deep Learning technologies
- Interest in Machine Learning approaches
- Interest in Semantic Web technologies

Contact person:

Russa Biswas

rusa.biswas@kit.edu

rusa.biswas@fiz-karlsruhe.de