Information Extraction Seminar Topics

Dario Stojanovski

Ludwig Maximilian University of Munich

stojanovski@cis.lmu.de

October 21, 2019
Reminder: these topics should all be in English
Self-training for Temporal Relation Extraction

- Matching events to specific times. Important in the medical domain for tracking progression of illnesses.

- Challenging to obtain large labeled datasets. How to handle scenarios with limited labeled instances?

- Explain how self-training works for this task. Contrast to other approaches to semi-supervised learning.

- Analysis of improvements and errors.

Sources and possible papers

Cross-lingual Named Entity Recognition

- Motivation for cross-lingual NER.
- How is cross-lingual NER different than monolingual NER?
- What are the potential benefits?
- Explain what kind of data is needed and how to use it?
- Explain the method and features used in the paper.

Sources and possible papers

Automatically Labeled Data for Event Extraction

- Overview of hand-labeled data for event extraction. What problems do we face when using this data?
- Can automatically labeled data solve them?
- Explain the process of generating automatically labeled data.
- Present the method for event extraction.

Sources and possible papers

- Chen, Yubo, et al. ”Automatically Labeled Data Generation for Large Scale Event Extraction.” Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics
Aspect-based Sentiment Analysis

- Motivation for aspect-based sentiment analysis.
- How is it different than traditional sentiment analysis?
- How does the data look like? Is it challenging to create?
- Neural networks - what architecture is best suited for the task?

Sources and possible papers

Hyperpartisan News Detection

- Fake vs Hyperpartisan news - differences and similarities.
- Present the shared task and explain how is the task defined.
- What is the data? How was it created, annotated etc. How were the teams evaluated?
- What are the shared task conclusions?
- Present the basic approach of the best team.

Sources and possible papers