Information Extraction Seminar Topics

Simon Rieß
riess.simon@yahoo.de

21.10.2019
My topics can be done in German or English
Hate Speech detection

▶ Offensive language in social media
▶ How to detect hate speech automatically?
▶ What data sets are available?
▶ How to collect and augment such a data set?
▶ How to build counter-narratives?

Sources:

▶ Punyajoy Saha, Binny Mathew, Pawan Goyal, Animesh Mukherjee, ”HateMonitors: Language Agnostic Abuse Detection in Social Media”, Arxiv 2019
▶ Y.L. Chung, E. Kuzmenko, S.S. Tekiroglu, M. Guerini, ”CONAN - COunter NAarratives through Nichesourcing: a Multilingual Dataset of Responses to Fight Online Hate Speech”, ACL 2019
Propaganda detection

- Different propaganda techniques
- How to collect relevant data and augmentation?
- How to evaluate?
- How to build models that detect various kinds of propaganda?

Sources:
- Giovanni Da San Martino, Seunghak Yu, Alberto Barrón-Cedeño, Rostislav Petrov, Preslav Nakov, ”Fine-Grained Analysis of Propaganda in News Articles”, EMNLP 2019
Fake News detection

- Available data sets?
- Describe models dealing with this problem
- How do they differ?
- What is BERT?

Sources:
- Álvaro Ibrain Rodríguez, Lara Lloret Iglesias, ”Fake News Detection using Deep Learning”, Arxiv 2019
Question Answer Systems

- What are the main components of a QA system?
- Describe the workflow in a QA system
- Give an overview of important models
- How to evaluate Question Answer systems?

Sources:
Question Generation

- How to ask relevant questions?
- How to include relevant key words?
- What are Question-Answer systems?
- How to use reranking to improve performance?

Sources:
- Jiazuo Qiu, Deyi Xiong, ”Generating Highly Relevant Questions”, EMNLP 2019