Statistical Consulting

Topic-specific sentiment analysis for tweets by German MPs

Department of Statistics
Ludwig-Maximilians-Universität München

New data sources offer vast potential for studying public opinion in the political environment. Frequently arising research questions concern the nature of topics discussed and the sentiment expressed toward them. We perform sentiment analysis for tweets by German MPs issued since the last federal election in September 2017, scraped from the Web, with special regard to topical context. The contribution of this work is twofold. First, we explore how such analysis can be conducted with either standard machine learning approaches for tabular data or more complex, BERT-based deep learning solutions. Our results suggest that casting the problem as a standard machine learning task achieves acceptable predictive performance and requires extensive feature engineering. BERT-based models clearly outperform this simplistic approach while avoiding the need for hand-crafted inputs. The ability to draw on knowledge acquired during the pre-training phase possibly supports the Transformers in solving the difficult task of classifying relatively few observations in a sparsely populated feature space. Post-training on the large pool of unlabeled data typically available in such settings is also helpful. We further note that topic modeling does not aid classification in this application, in fact complicating it, but attribute this largely to the fact that extracting topics from this kind of data is challenging even for the human annotator. Second, we provide teaching material for instructing fellow researchers in the applied techniques. With this collection, which will remain publicly available, we hope to facilitate the entrance to NLP for practitioners from other disciplines.

Speakers Asmik Nalmpatian & Lisa Wimmer

Project partner Prof. Dr. Paul Thurner
Department of Political Science

Supervisors Prof. Dr. Christian Heumann
Matthias Aßenmacher
Department of Statistics

Date Monday, July 12th, 2021, 16:00 s.t.

Meeting ID 918 5472 5495, passcode 413775

`https://lmu-munich.zoom.us/j/91854725495?pwd=NHVLemt4ZzZHNk8vSUpHWHBZamZjUT09`