Predicting Political Parties from Text

Mattia Masiero
Matrikelnummer 6100692
mattia.masiero@student.uni-tuebingen.de

Tim Weiland
Matrikelnummer 6010188
tim.weiland@student.uni-tuebingen.de

Abstract

We are planning to use a collection of Bundestag speeches held since 2017 to see how well the association to a political party can be predicted from text using simple features such as PCA applied to bag-of-words, text/sentence length or width of vocabulary. We plan on using logistic regression for the classification. To analyze if and how speeches have changed over time, we plan on combining the OpenParliament dataset with another dataset such as ParlSpeech or the German Political Speeches Corpus. If time permits, we would like to use the results of our analysis to create a political slogan generator.