Word embeddings for predicting political affiliation based on twitter data

IBRAHIM ABDELAZIZ[1], OLIVER BERG[1], ANGJELA DAVITKOVA[1], MD ASHRAF HOSSAIN[1], CHARU JAMES[1], SHRIRAM SELVAKUMAR[1], KUMAR SHRIDHAR[1], and SAURABH VARSHNEYA[1]

[1]Technische Universität Kaiserslautern

November 19, 2017

1 Investigated Problem

The modern world of social media knows a plethora of means to communicate ones personal opinion and political alignment. With the platform *Twitter*, figures of political interest are expressing their standpoints in small-sized 144-character texts, which contain a comprised message specific to the general public. This yields great potential for automated analysis of party affiliations to classify political persons of interest within the political spectrum.

The aim of this project is to predict the political affiliation of German *Twitter* users based on the content of their tweets, this may potentially be used in various applications such as measuring the percentage of support for political parties on the social media, or predicating the results of elections based on data acquired from *Twitter*.]

2 Literature Research

A thorough introduction to political affiliation analysis was provided in *"Predicting political party affiliation from text" by Biessmann et. al.*, where political motives were shown to be consistently predictable with an accuracy better than chance.

[.. papers for methods deployed in example paper and/or our later approaches ..]

3 Considerations towards Data

According to the presented task description of analyzing political affiliation based on Twitter-data, the final comparison occurs on the basis of **tweets** made by political figures on the platform Twitter.

Along the ground work established in *"Predicting political party affiliation from text" by Biessmann et. al.*, the learning models are also deployed against **parliament discussion data** as well as **party manifesto data** to establish a categorical groundwork.

[.. more specifically: which data should be used how ..]

[...]

4 Deployed Methods

[.. which algorithms / categories of algorithms to test first/second/third ..]
[.. which approaches may yield which types of results; generally: what is to be expected? ..]

5 Work Packages

[–we should divide the workoload into number of packages, and assign group of people and estimated time or deadline to each one]