Research report - Hyperpartisian NLP

Florian Biebel and Daan Middendorp

May 2019

1 Introduction

This report is part of a project at the Technische Universität Berlin, called Advanced Projects at the Quality and Usability Lab. During this project, the problem of hyperpartisian news detection will be solved using natural language processing. Hyperpartisian news is news that is extremely biased or sharply polarized in favor of a political party.

This problem comes from a task in the Semantic Evaluation[1] competition which is held simultaneously during this project. This makes it possible to use the dataset provided to the participants of the competition. The goal of this project is to achieve a deeper understanding of the inner working of such a system and learn how to build such a system ourselves.

2 Existing approaches

2.1 Preprocessing

2.2 LSTM

Long short-term memory a form of an artificial neural network. The difference with a recurrent neural network is that LSTM networks are able to use information from the past to understand the input. E.g. a video frame can be better understood if the previous video frames are also taken into account. This is also valueable for text processing, because hyperpartisian news detection requires a deeper understanding of the link between words.

---

[1] https://pan.webis.de/semeval19/semeval19-web/