

A poor man's opinion detection tool: training a model with a handful of data

Master/Bachelor Thesis

Motivation

Opinion detection aims to detect an author's view towards a certain topic and has become a key component in everyday applications such as fake news detection and argumentation. While state-of-the-art deep learning models are ideal for this problem, they need lots of labeled training data that is often expensive to obtain. In order to reduce data collection costs, we can use few-shot-learning models which are specifically designed for learning with a very small amount of training data. Since most of our data comes from social media, we need a tool that can differentiate useful text from noisy, unopinionated posts. To this end, this thesis aims to develop an opinion detection model that can efficiently differentiate opinionated from non-opinionated text and is trained only on a small amount of data.

Difficulty

Analysis



Programming



Literature



Task Description

- Processing literature about state-of-the-art few-shot learning classification models
- Training different variations of the opinion detection model:
 1. evaluate the model on a large number of completely new topics (zero-shot opinion detection) and
 2. evaluate the model on a large number of topics with very few training examples (few-shot opinion detection)
- Performing an error analysis

Target: Abortion

Opinionated: Yes



Forcing a woman to carry the child of her rapist is cruel. You have to think about the psychological implications

Target: Abortion

Opinionated: No

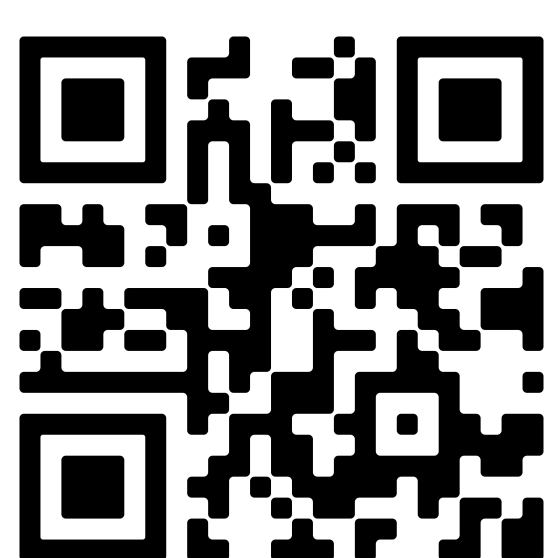


Agreed. The reality is they're just used to annoy the other side. I don't see the utility in that.

Contact

Prof. Dr. Lucie Flek &
Dr. Flora Sakketou

✉ sakketou@mathematik.uni-marburg.de



References

- [1] Yaqing Wang and Quanming Yao. Few-shot learning: A survey. *CoRR*, abs/1904.05046, 2019.
- [2] Emily Allaway and Kathleen R. McKeown. Zero-shot stance detection: A dataset and model using generalized topic representations. *CoRR*, abs/2010.03640, 2020.
- [3] Rui Liu, Zheng Lin, Yutong Tan, and Weiping Wang. Enhancing zero-shot and few-shot stance detection with commonsense knowledge graph. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 3152–3157, Online, August 2021. Association for Computational Linguistics.