

Few-Shot Learning for Argument Aspects of the Nuclear Energy Debate

Lena Jurkschat, Gregor Wiedemann,
Maximilian Heinrich, Mattes Ruckdeschel, Sunna Torge



Diese Maßnahme wird gefördert durch die Bundesregierung aufgrund eines Beschlusses des Deutschen Bundestages. Diese Maßnahme wird mitfinanziert durch Steuermittel auf der Grundlage des von den Abgeordneten des Sächsischen Landtags beschlossenen Haushaltes.



Motivation

- Argument Mining application in social/communication science is rare
- Argument Mining mainly focussed on structural argument components

Problems

1. Absence of semantic argument components
2. Label effort for supervised methods



Supervised aspect based argument mining (ABAM) for field of ‚nuclear energy‘

Supervised Aspect-based Argument Mining (ABAM)

Main Contributions:

1. Build the **Argument Aspect Corpus – Nuclear Energy** containing english sentences with content describing annotations
2. Trained transformer based models in a few-shot setting for aspect classification
3. Case study: Apply the best classifiers on arguments from newspaper articles



The AAC-NE Corpus

- Arguments **from UKP SAMCorpus** [0]
 - Topic „Nuclear Energy“
 - 740 pro and 564 con arguments

“Nuclear plants require significant quantities of water to remove the waste heat - usually into the environment.” → CON

- Hierarchical **Clustering** using S-BERT embeddings
 - Semantic groups/subtopic detection
 - Grouping into 15 Clusters

[0] <https://aclanthology.org/D18-1402/>

The AAC-NE Corpus

Definition Aspects:

- Aspect defined as subtopic of field
- 9 aspects from 15 clusters

Annotation via crowd-sourcing

- Tool: Doccano (open source)
- Trained student annotators
- Majority vote label decision
- Krippendorff- α for Inter-coder-reliability

Aspect	Train	Val	Test	Σ	α
alternatives	100	16	21	137	0.69
costs	98	17	29	144	0.72
environment	209	27	64	300	0.74
innovation	33	2	8	43	0.38
reactor safety	112	17	43	172	0.59
reliability	47	5	10	62	0.36
waste	87	5	26	118	0.80
weapons	52	11	15	78	0.77
other	120	23	29	172	0.49
all	858	123	245	1226	0.62

Aspect Classification – Fine-tuning Transformers

1. Find baseline: Fine-tuned transformer model

Model	Precision		Recall		F1		Accuracy	
	mean	std	mean	std	mean	std	mean	std
albert-base-v2	68.5	±8.6	67.8	±5.2	67.2	±5.7	75.1	±2.0
bert-base-cased	72.2	±3.0	75.3	±3.2	71.6	±3.3	77.6	±3.4
bert-large-cased	67.1	±2.7	70.7	±4.2	67.0	±3.1	75.3	±1.7
electra-large	72.0	±2.6	73.8	±2.9	70.0	±1.7	73.8	±2.7
roberta-large	75.6	±4.7	77.2	±2.2	74.1	±2.5	78.4	±2.9
xlm-roberta-large	69.3	±3.7	69.5	±5.4	67.5	±3.6	73.5	±1.5

baseline model

2. Few-shot training

- Train with $n \in \{1,2,4,8,10,50,100\}$ samples for each label
- Five training repetitions
- Compare Few-Shot models with baseline model

Aspect Classification – Models

Models:

- TARS - Task aware representations for multi label text classification

“Firstly If Australia does not use nuclear energy Australia would let off greenhouse gasses. [SEP] costs”

- PET – Pattern Exploiting Training using Pattern Verbalizer Pairs (PVP’s)

Pattern: “This argument addresses the aspect [label] in the nuclear energy debate. [SEP] [text]”

Verbalizer: aspects verbalized by single tokens

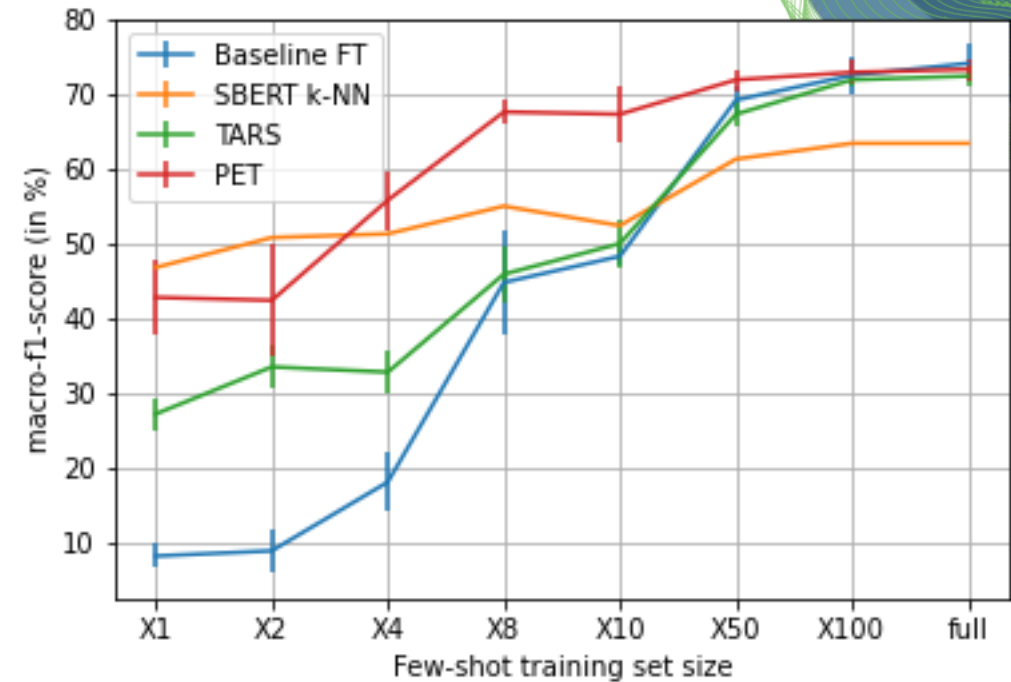
- SBERT *k*-NN

- *all-mpnet-base-v2* of Sentence-BERT → similarity of argument and *label sentence*

Label Sentence: “In the nuclear energy debate, the argument is about [label]”

Results

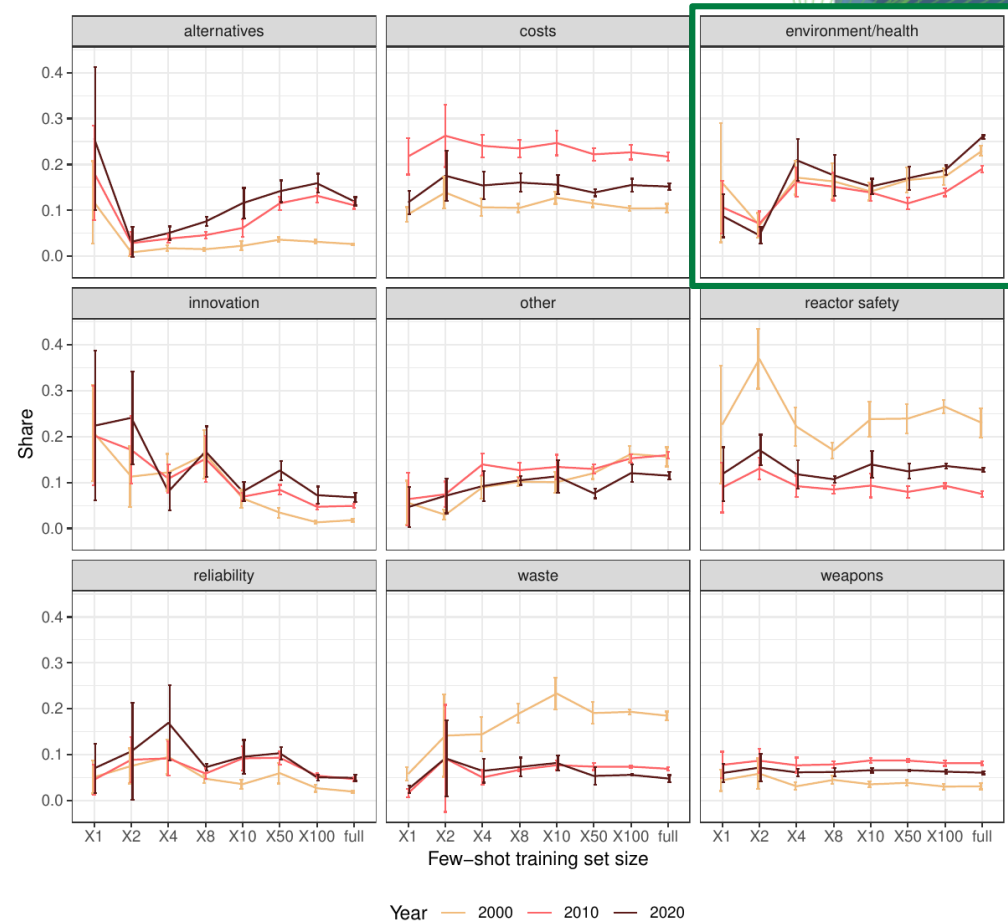
- Baseline transformer model fails on small datasets
- SBERT k-NN doesn't benefit from larger training sets
- PET model most stable results
 - >67% f1-score for min. 8 samples per category



Application on Newspaper Articles

„The Guardian“ Articles [0]

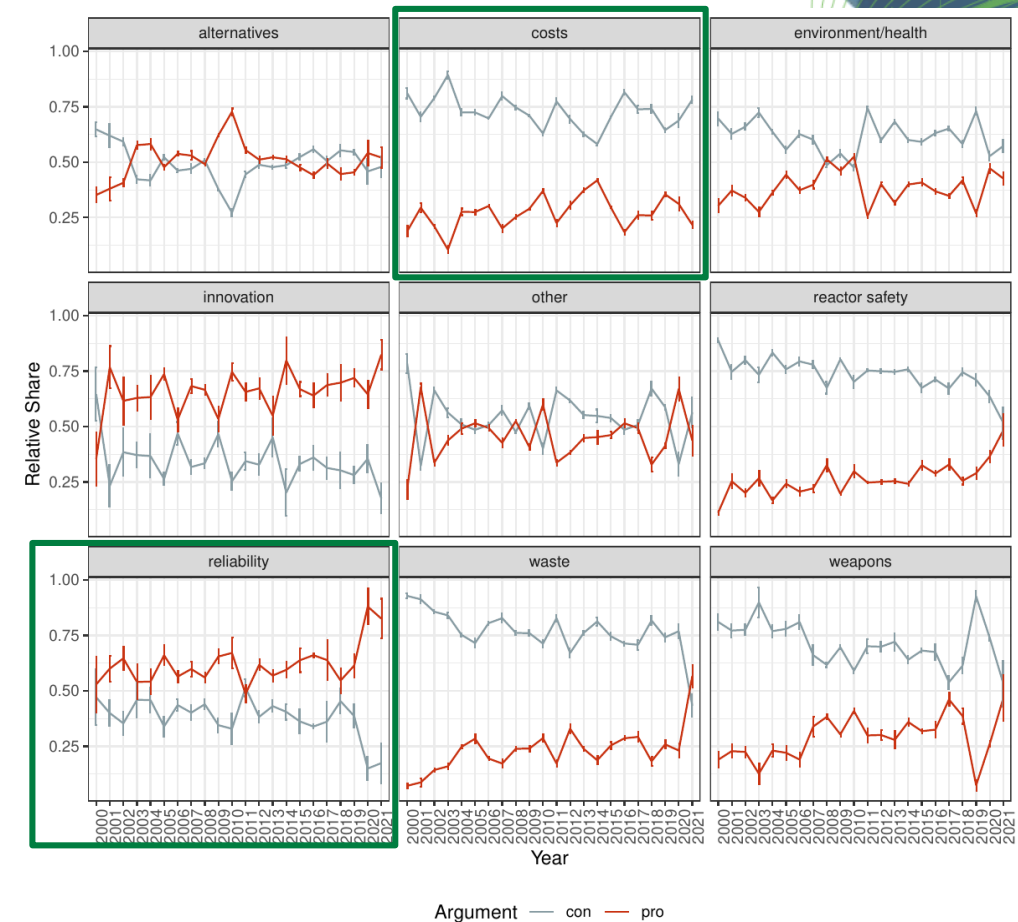
- 3268 „environment/nuclearpower “ articles
- 26.571 argumentative sentences classified
 - 16.902 con
 - 9.669 pro
- Evaluation: stability of prediction for different training set sizes



[0] <https://open-platform.theguardian.com>

Pro and Contra Arguments over Time

- Pro and con relative shares from 2000-2021
- Indicate discourse shifts over time
 - „Costs“ stable contra
 - „reliability“ makes up to pro majority



Conclusion

- Introduction of
 - *supervised* aspect based argument mining
 - AAC-NE dataset
- Detection, definition and labeling of aspects is a labour intensive procedure
- Annotation effort for stable classification manageable through few-shot learning techniques
 - 50-100 annotations per category minimum
- Application: determination of clear shifts in public discourse for benefit of nuclear energy

Questions

{lena.jurkschat, sunna.torge}

{g.wiedemann, m.ruckdeschel}

mheinrich

@tu-dresden.de

@leipzig-hbi.de

@informatik.uni-leipzig.de