Searching Arguments in German with ArgumenText

Chris Stahlhut
Research Training Group KRITIS
Ubiquitous Knowledge Processing Lab
Darmstadt, Germany
www.ukp.tu-darmstadt.de

1 INTRODUCTION

Searching for arguments in large collections of documents is an important task in decision making. Although current information retrieval systems are working well in finding documents and passages relevant to a query, they do not provide specialised support for finding arguments. Finding arguments requires to distinguish argumentative from non-argumentative passages and is one of the core tasks in Argument Mining (AM), e.g. detecting argumentative sentences in heterogeneous sources.

The ArgumenText project [2] aims at building an argument search engine for heterogeneous sources[1]. The current system considers a sentence an argument if the sentence is "expressing evidence or reasoning that can be used to either support or oppose a given topic" [3]. When a user searches for arguments on a topic such as "nuclear energy", it first retrieves relevant documents via Elasticsearch from a large collection of documents, such as Common Crawl[2]. In the second step, it detects the arguments related to the same topic using a birectional Long-Short Term Memory (BiLSTM) with attention [3]. Afterwards, it detects the stance towards the topic and presents the arguments with the highest score to the user. However, as with most AM methods, the current system is only available in English.

2 ARGUMENT MINING IN GERMAN

In extending the language capabilities of an argument search engine, e.g. to enable searching for arguments in German, we are faced with the challenge of creating AM models in other languages. Before we can train an AM model in German, we need to create training data. This data can be created by manually labelling German sentences as not-argumentative, pro-, or con-argument. Alternatively, we can translate an existing corpus to German and keep the labels. When done manually, both approaches are labour intense, expensive, and require a large amount of time. In the case of translating the training data, however, machine translation systems have reached a sufficient quality to automate this step [1]. We therefore decided not to repeat the process of manually labelling sentences but instead use the Google Translation API to translate our existing corpus [3] from English to German.

After translating the corpus, we repeated the experiments of Stab et al. [3] with the same architecture, hyper-parameter settings, and randomisation seeds to evaluate how well the German models perform in an in- and cross-topic setup. Table 1 shows that the translation did not affect the performance negatively in the cross-topic setup; the in-topic scores dropped to a comparable level.

Table 1: The macro F1 scores are averaged across topics.

	in-topic		cross-topic	
	DE	**EN**	**DE**	**EN**
BiLSTM	0.636	0.721	**0.626**	0.592
BiLSTM+cos	**0.646**	0.732	0.624	0.626
BiLSTM-att	0.633	0.741	0.617	0.623
BiLSTM-att+cos	0.621	0.736	0.619	0.658

Our next steps are evaluations of additional machine learning techniques, e.g. adversarial learning for better AM, and an extrinsic evaluation of the quality of the predictions with German speaking users. Furthermore, the relation to decision making offers many additional applications in social science research. For example, political scientists can use such an argument search engine to improve the understanding of controversial topics and what led to a specific decision; or, historians can use it to analyse the discourse around important historical decisions. Supporting these examples comes with additional challenges, such as (1) *complex queries*, in political and historical research, queries are often much more complex than "nuclear energy" and (2) *concept drift*, a researcher might decide the focus on specific aspect of a topic and thereby implicitly change the requirements. We aim to address these challenges by interactively improving the AM model. Finally, we can use the adaption to the German data and users as a test bed for other languages.

ACKNOWLEDGMENTS

This work has been supported by the German Research Foundation (DFG) as part of the Research Training Group KRITIS No. GRK 2222/1, by the German Federal Ministry of Education and Research (BMBF) under the promotional reference 03VP02540 (ArgumenText), and by the German Federal Ministry of Education and Research under the promotional reference 01UG1816B (CEDIFOR).

REFERENCES

[1] Steffen Eger, Johannes Daxenberger, Christian Stab, and Iryna Gurevych. 2018. Cross-Lingual Argumentation Mining: Machine Translation (and a Bit of Projection) Is All You Need!. In *Proceedings of the 27th International Conference on Computational Linguistics (COLING 2018)*. to appear.

[2] Christian Stab, Johannes Daxenberger, Chris Stahlhut, Tristan Miller, Benjamin Schiller, Christopher Tauchmann, Steffen Eger, and Iryna Gurevych. 2018. ArgumenText: Searching for Arguments in Heterogeneous Sources. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*. Association for Computational Linguistics, New Orleans, Louisiana, 21–25.

[3] Christian Stab, Tristan Miller, and Iryna Gurevych. 2018. Cross-Topic Argument Mining from Heterogeneous Sources Using Attention-Based Neural Networks. *arXiv:1802.05758 [cs]* (Feb. 2018). arXiv:cs/1802.05758

[1] A demonstrator is publicly available at https://www.argumentsearch.com/
[2] https://commoncrawl.org/

DESIRES 2018, August 2018, Bertinoro, Italy