Open Domain Rhetorical Question Detection by Representation Learning via Soft-Prompting

Josef Baloun[1], Jiří Martínek[2], Christophe Cerisara[3], Pavel Král[4]

1 Introduction

Automatic detection of rhetorical questions is essential for many real applications, including chatbots or social network information retrieval – Paul et al. (2011) found out that 42% of all questions on English Twitter are rhetorical. However, this task is very challenging even for large pre-trained language models because a rhetorical question is very similar to a common one, and moreover, only a few annotated corpora exist. Therefore, we propose a novel approach using soft-prompting to train the model for sentence representation in order to deal with automatic rhetorical question detection with only a few annotated data. Further, we focus on several datasets, their relations and usefulness for each other given a specific model.

2 Representation Learning via Soft-Prompting

In the case of fine-tuning, the entire pre-trained language model is adjusted. This is problematic especially when working with little data, as it usually results in losing the pre-trained knowledge. Soft-prompting is more suitable for such a problem because it does not modify the pre-trained model. Instead, it learns several prompts (see Fig. 1) that can change the behaviour of the model or extract the desired information.

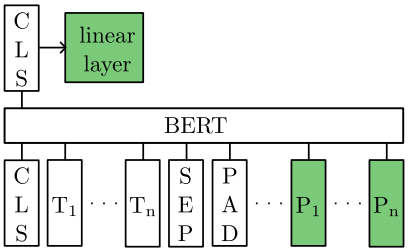

Figure 1: Soft-prompt model, trainable in green

The representations are then learned using triplet loss in a contrastive training manner. Finally, the representation is used as input for a conventional classifier (e.g. K-NN).

3 Dataset alignment

We focused on several datasets including SwDA (Godfrey et al., 1992), RQueT (Kalouli et al., 2021) and SarcV2 (Oraby et al., 2017). They relate to rhetorical questions, information

[1] student of the doctoral degree program Applied Sciences, field of study Computer Science, e-mail: balounj@kiv.zcu.cz
[2] DCSE, University of West Bohemia, e-mail: jimar@kiv.zcu.cz
[3] Université de Lorraine CNRS, LORIA, F-54000, Nancy, France, e-mail: cerisara@loria.fr
[4] DCSE, University of West Bohemia, e-mail: pkral@kiv.zcu.cz

seeking questions and sarcasm, respectively. Intuitively, these classes should have something in common. Once we have representations, we can visualize them and compare the alignment between datasets. The difference between RQueT and SarcV2 alignment with SwDA in Fig. 2 is striking. We can conclude that RQueT is useful for the target task, whereas SarcV2 is not because the learned representations does not distinguish target SwDA samples.

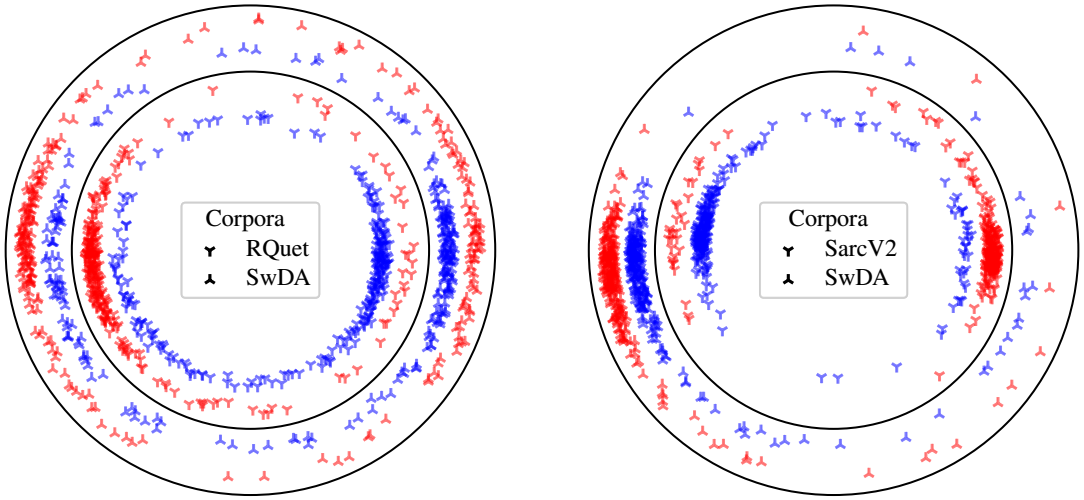

Figure 2: Representations trained on RQueT (left) and SarcV2 (right)

4 Results

Using the alignment information, we further adapted the model for several datasets at once and obtained 75 % accuracy in a real 10-shot setup on SwDA. State-of-the-art using the whole dataset is 81 % Bhattasali et al. (2015).

Acknowledgement

This work was performed partly using HPC resources from GENCI–IDRIS (Grant 2023-AD011011668) and funding from the project LUE OLKi and from grant no. SGS-2022-016 Advanced methods of data processing and analysis.

References

Bhattasali, S., Cytryn, J., Feldman, E., and Park, J. (2015). Automatic identification of rhetorical questions. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 743–749.

Godfrey, J. J., Holliman, E. C., and McDaniel, J. (1992). Switchboard: Telephone speech corpus for research and development. In *Proceedings of the 1992 IEEE International Conference on Acoustics, Speech and Signal Processing - Volume 1*, ICASSP'92, page 517–520, USA. IEEE Computer Society.

Kalouli, A.-L., Kehlbeck, R., Sevastjanova, R., Deussen, O., Keim, D. A., and Butt, M. (2021). Is that really a question?: Going beyond factoid questions in nlp. In *14th International Conference on Computational Semantics: IWCS 2021*, pages 132–143.

Oraby, S., Harrison, V., Reed, L., Hernandez, E., Riloff, E., and Walker, M. (2017). Creating and characterizing a diverse corpus of sarcasm in dialogue. *arXiv preprint arXiv:1709.05404*.

Paul, S. A., Hong, L., and Chi, E. H. (2011). What is a question? crowdsourcing tweet categorization. *CHI 2011*.