Cross-Lingual Metaphor Detection for Low-Resource Languages

Anna Hülsing[1], Sabine Schulte im Walde[2]
[1]Universität Hildesheim, [2]Universität Stuttgart
anna.huelsing@uni-hildesheim.de, schulte@ims.uni-stuttgart.de

Research on automatic metaphor detection in a multilingual setup has gained momentum only recently. When using multilingual pretrained models (PLMs), it is however unclear how the amount of target language data used for pretraining affects the performance of the model. Especially for languages where little data is available, PLMs often struggle, and non-neural models might be more suitable. We therefore compare neural and non-neural cross-lingual models for Russian, German and Latin as target languages, because the amount of data used to pretrain our multilingual PLM (mBERT, Devlin et al. 2019) varies greatly across these three languages. English is used as the source language.

Our metaphor detection focuses on word-based classification, as in the following example from the metaphor dataset by Köper and Schulte im Walde (2016). We define a binary classification task to detect whether the underlined target word is used metaphorically in the given context, or not.

(1) Über die Zeit hatte sich in ihnen Sehnsucht und Verlangen aufgestaut. (translation: "Over time, longing and desire had dammed up inside them.") → metaphorical

In order to find the best-performing model for each target language, we firstly carry out zero-shot classification with mBERT, where we fine-tune the PLM to metaphor detection on source language data. Secondly, we apply few-shot classification with mBERT, where we first fine-tune our PLM on source language data, and then fine-tune again on 20 target language sentences. Thirdly, we use the adapter architecture MAD-X (Pfeiffer et al. 2020), where small amounts of trainable weights (the adapters) are injected into the PLM such that they improve zero-shot classification. Lastly, as our non-neural model we use a random forest classifier that utilizes a vector space model and conceptual features (abstractness and supersenses) – similarly to the model introduced by Tsvetkov et al. (2014).

References: • Devlin, J. et al. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *NAACL*, 4171-4186. • Köper, M. & S. Schulte im Walde (2016). Distinguishing literal and non-literal usage of German particle verbs. In *NAACL*, 353-362. • Pfeiffer, J. et al. (2020). MAD-X: An adapter-based framework for multi-task cross-lingual transfer. In *EMNLP*, 7654–7673. • Tsvetkov, Y. et al. (2014). Metaphor detection with cross-lingual model transfer. In *ACL*, 248–258.