Few-Shot Cross-Lingual Stance Detection with Sentiment-Based Pre-Training

Momchil Hardalov[1,2] Arnav Arora[1,3] Preslav Nakov[1,4] Isabelle Augenstein[1,3]

[1]Checkstep Research

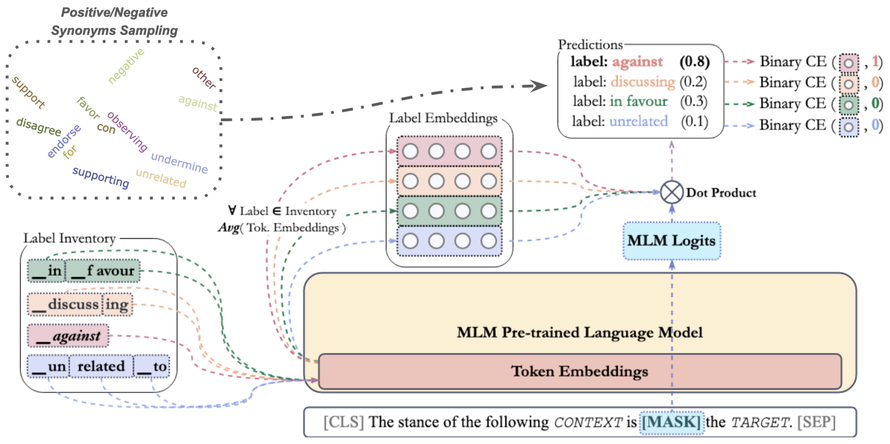
[2]Sofia University [3]University of Copenhagen [4]Qatar Computing Research Institute, HBKU

{momchil, arnav, preslav.nakov, isabelle}@checkstep.com

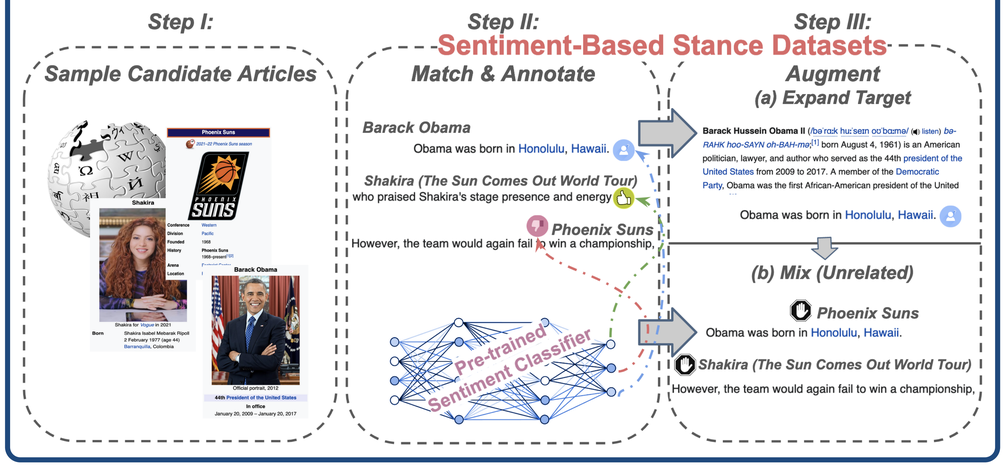
Summary

- Cross-lingual stance datasets:
  - **100,000** examples (60K for training)
  - mostly **small in size** (< 2K training examples)
  - use **different labels**, e.g., *Agree*, *Support*, *Argument for*, etc.
  - hard to use in cross-learning studies
- We explore cross-lingual few-shot learning (**PET with a label encoder)**.
- We generate (noisy) **training data with self-supervised labelling** using sentiment analysis as a proxy.
- We experiment with knowledge transfer:
  - transfer learning from English stance
  - multi-dataset learning (MDL) in a cross-lingual few-shot setting
- We achieve **strong performance** in both few-shot and full-resource scenarios.
- We create a **challenging testbed for cross-language evaluation.**
- Code & Data:

Pattern-Exploiting Training (PET) with a Label Encoder



Sentiment-Based Stance Datasets

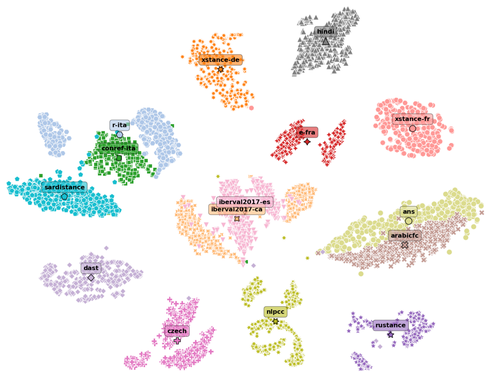


Contributions

- The largest study of cross-lingual stance detection to date
- Exploring pattern-exploiting training (PET) for cross-lingual learning
- Novel label encoding mechanism
- Novel semi-supervised approach to produce automatically labelled instances
- Promising results in a cross-lingual testbed.

Cross-Lingual Datasets

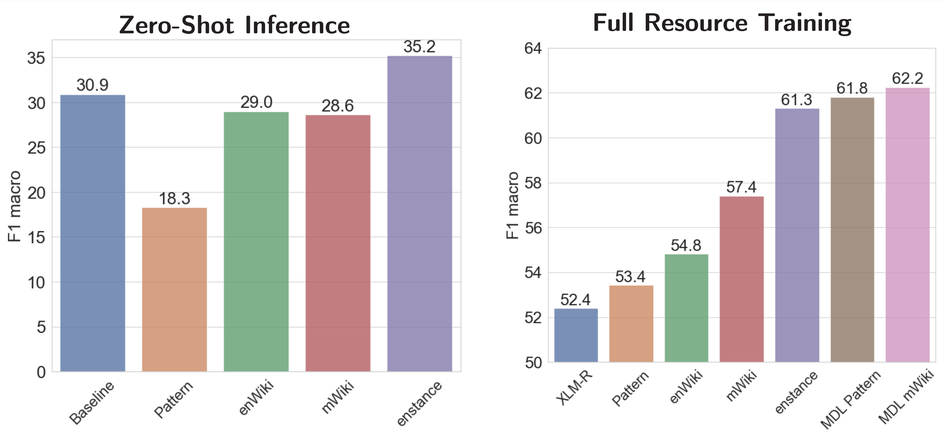
- **15 diverse datasets:**
  - 12 languages
  - 6 language families
  - yet, mostly small in size
- Different **topics**
  - News, Politics, Debates, etc.
- Over **30 unique labels**
  - (well-known) *Favor*, *Against*, etc.
  - (additional) *Neutral*, etc.
  - (unusual) *Other*, *Question*, etc.



Pre-training Datasets

- Sentiment-based stance (**en/mWiki**)
  - Sampled from **12,000 Wiki articles**
    - multilingual: 1K per-language
    - English: all English
  - **300,000** examples (8:1:1 split)
- English stance (**enstance**)
  - **16** datasets
  - 24 unique labels
  - **250,000** examples (154K for training)

Training with All/No Data



Few-Shot Results

- Training on 32 examples yields **large variance** and **instability** of the model
  - σ varies from 1.1 to 8.9 (3.5 avg.)
  - en/mWiki reduce avg. σ to 2.7
  - MDL drops σ to under 1.7
- **Multilingual sentiment** pre-training is **better** than **English sentiment** pre-training.
- **English stance** adds **3–12 F1 points** in a few-shot setting compared to other models.
- There is a **risk** of introducing **additional bias with the pre-training**.