Zero-Shot Cross-Lilingual Multi-target Text Stance Detection Based on Pre-trained Models

狮驼岭队

艾子豪　李晨　石涛　胡元泉　李郁青　王姣姣

Outline

- Problem Definition
- Existing Approches
- Limitations with Existing Approaches
- Proposed Methodology
- X-stance Dataset
- Experiments and Results
- Future Work

Problem Definition

- Text stance detection aims to determine the position of a person towards a target (a concept, idea, event, etc.) from a piece of text he/she produces. Available stances are: {Favor, Against, Neutral}.

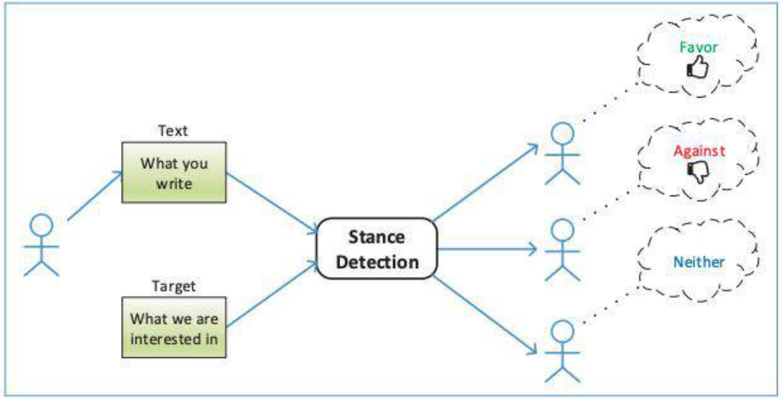

Figure 1: An illustration of text stance detection.

Existing Approaches

- **Traditional machine learning approaches**: support vector machine, decision trees, naïve bayes

- **Ensemble Learning approches**: majority voting, proprietary ensemble learners

- **Deep Learning approaches**: CNN, RNN, large scale pre-trained models

Limitations with Existing Approaches - Limited Multilingual Resources

- Most existing research in stance detection has been limited to work with a single language, with little work on cross-lingual stance detection, as the multilingual datasets available today are scarce and relatively small[1]

- While English datasets exist for various domains and in different sizes, non-English and multilingual datasets are often small and focus on narrow, potentially country or culture-specific topics[2]

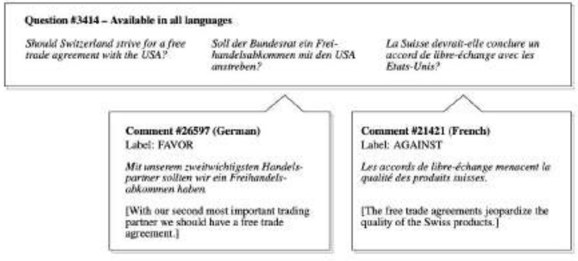

Figure 2: An illustration of multilingual stance detection.

[1] Joshi, P.; Santy, S.; Budhiraja, A.; Bali, K.; and Choudhury, M. 2020. The State and Fate of Linguistic Diversity and Inclusion in the NLP World. In ACL, 6282–6293. Online.
[2] Lozhnikov, N.; Derczynski, L.; and Mazzara, M. 2020. Stance Prediction for Russian: Data and Analysis. In SEDA, 176–186.

Limitations with Existing Approaches - Multil-Target Scenarios

- Most research on stance detection treat different target entities separately (i.e., single-target stance detection) and ignore the underlying relationship among targets[3], which is complex to model
- Existing multi-target stance detection focused on a per-target-pair training strategy[4], which is inefficent and time-consuming

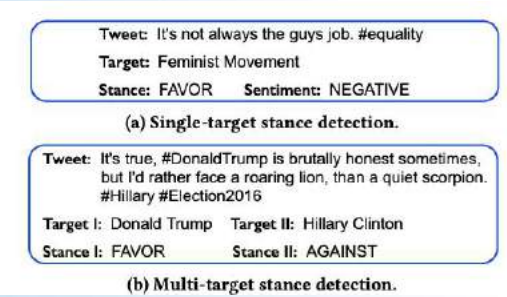

Figure 3: An illustration of multi-target stance detection where target entities are closely related.

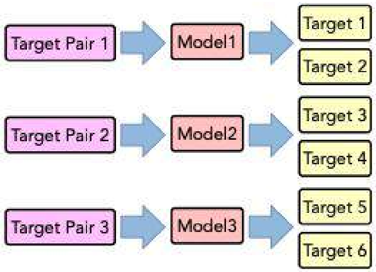

Figure 4: An illistration of previous work on multi-target stance detection, which adpted a per-target-pair training strategy.

[3] Chenguang Chen, Wen Xi, and Bin Zhou. 2021. Multi-Target Stance Detection with Multi-Task Learning. In Proceedings of the 2020 9th International Conference on Computing and Pattern Recognition (ICCPR 2020). Association for Computing Machinery, New York, NY, USA, 111–116. https://doi.org/10.1145/3436369.3436473
[4] Li, Yingjie, and Cornelia Caragea. "A multi-task learning framework for multi-target stance detection." In Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021, pp. 2320-2326. 2021.

Proposed Methodology

Training Phase

- Use pre-trained cross-lingual **XLM-RoBERTa (XLM-R)**[5] which has been pre-trained jointly in 100 languages as our model and finetune it on the **Multilingual X-stance dataset**[6]

- Interpret the X-stance task as sequence pair classification and designate the question(text) as segment A and the comment as segment B

Inference Phase

- Use XLM-R to perform **Name Entity Recognition(NER)** and extract all the target aspects in the data and meanwhile store the corresponding sentences

- Perform stance detection for every target and its corresponding sentence

[5] Conneau, Alexis, et al. "Unsupervised cross-lingual representation learning at scale." arXiv preprint arXiv:1911.02116 (2019).
[6] Vamvas, Jannis, and Rico Sennrich. "X-stance: A multilingual multi-target dataset for stance detection." arXiv preprint arXiv:2003.08385 (2020).

Proposed Methodology - Technical Advantage

- Excellent performance on multi-lingual dataset
- High NER accuracy for extracting arbitrary number of targets
- Accurate and efficient stance detection without explicitly modelling the structure of the sentences (interactions between each target word and opinion words)

X-stance Dataset

- A multilingual multi-target dataset which comprises 150 questions about different topics and 67k comments given by interviewees in Switzerland

Topic	Questions	Answers
Digitisation	2	1168
Economy	23	6899
Education	16	7639
Finances	15	3980
Foreign Policy	16	4393
Immigration	19	6270
Infrastructure & Environment	31	9590
Security	20	5193
Society	17	6275
Welfare	15	8508
Total (training topics)	**174**	**59 915**
Healthcare	11	4711
Political System	9	2645
Total (held-out topics)	**20**	**7356**

Table 1: The number of questions and answers per topic.

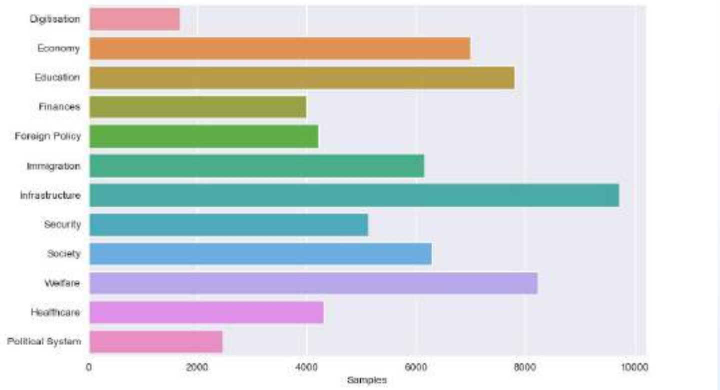

Figure 5: A visualization of the distributions of topics in X-stance dataset.

X-stance Dataset

- Questions are available in four languages: English, Swiss Standard German, French, and Italian

- We adopt the strategy of **No Italian and English samples** are seen during the training stage, making X-stance a case of **zero-shot cross-lingual transfer**

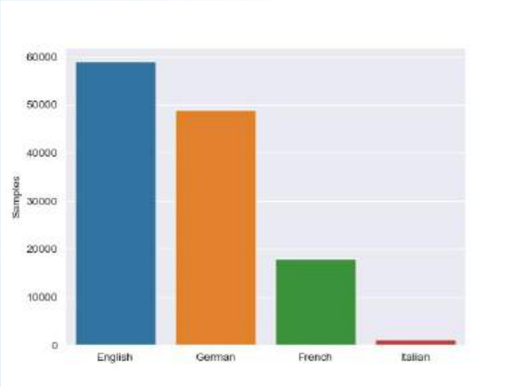

Figure 6: The distribution of questions in different languages.

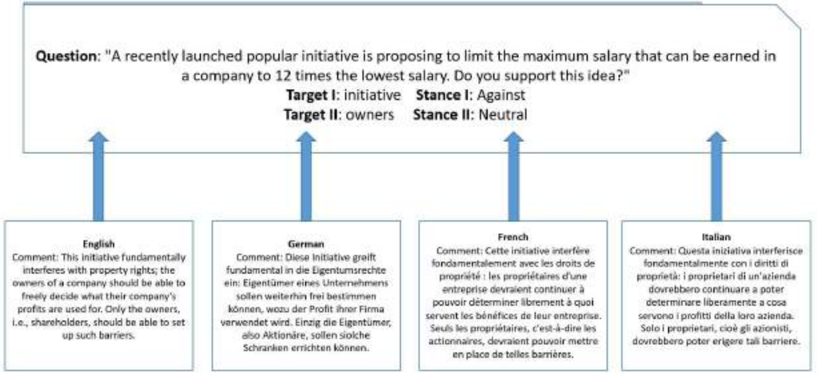

Figure 7: An example of a question and a comment in all four languages.

Experiments

- Remove all the English and Italian samples from the training set
- Use a batch size of 16 and a maximum sequence length of 512 subwords, and performed a grid search over the hyperparameters (learning rate and number of epochs) based on the validation accuracy
- Follow the standard recommendations for fine-tuning BERT: Adam with $\beta_1 = 0.9$ and $\beta_2 = 0.999$; an L_2 weight decay of 0.01; a learning rate warmup over the first 10% of the steps
- A Dropout layer with probability of 0.1 was set on all layers

Results

- XLM-R (zero-shot + NER) performs consistently better than existing baselines (majority class, fastText classifier, M-BERT) in most setttings

Model	EN F1-favor	EN F1-against	DE F1-favor	DE F1-against	FR F1-favor	FR F1-against	IT F1-favor	IT F1-against
Majority class (global)	35.3	34.8	33.4	32.9	34.7	35.0	34.2	34.6
Majority class (target-wise)	59.8	59.6	60.2	61.2	65.6	64.8	60.3	58.8
fastText	69.2	69.7	70.5	69.2	73.6	69.4	60.7	49.8
M-BERT	78.4	76.8	**77.2**	75.6	76.2	**77.0**	68.7	71.4
XLM-RoBERTa+NER (ours)	**82.3**	**79.2**	75.9	**76.4**	**76.3**	75.1	**70.4**	**71.6**

Table 2: The comparison of the performances of XLM-R (zero-shot + NER) and other existing approaches on the X-stance dataset.

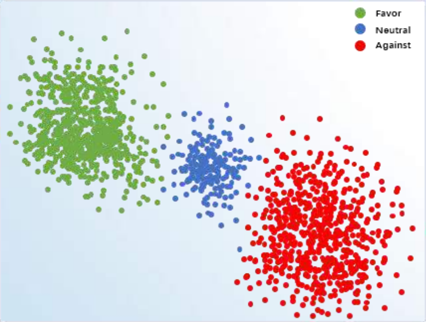

Figure 8: A visualization of the predicted stances using XLM-R (zero-shot + NER) on the X-stance test set.

Question: "A recently launched popular initiative is proposing to limit the maximum salary that can be earned in a company to 12 times the lowest salary. Do you support this idea?"

Comment: This initiative fundamentally interferes with property rights; the owners of a company should be able to freely decide what their company's profits are used for. Only the owners, i.e., shareholders, should be able to set up such barriers.

Target I: initiative
Predicted Stance I: Against
Ground-truth Stance I: Against

Target II: owners
Predicted Stance II: Neutral
Ground-truth Stance II: Neutral

Figure 9: An illustration of one English test sample in the X-stance dataset.

Results - Classification Error Analysis

- Some classification errors with extremely low confidences in ground-truth labels occur when the stances of these comments are expressed only on a very implicit level, or contain sarcasm and irony

Question	Comment	Gold Label	Prob.
Befürworten Sie eine vollständige Liberalisierung der Geschäftsöffnungszeiten? [Are you in favour of a complete liberalisation of business hours for shops?]	*Ausser Sonntag. Dies sollte ein Ruhetag bleiben können.* [Except Sunday. That should remain a day of rest.]	FAVOR	0.001
Soll die Schweiz innerhalb der nächsten vier Jahre EU-Beitrittsverhandlungen aufnehmen? [Should Switzerland embark on negotiations in the next four years to join the EU?]	*In den nächsten vier Jahren ist dies wohl unrealistisch.* [For the next four years this is probably unrealistic.]	FAVOR	0.005
Befürworten Sie einen Ausbau des Landschaftsschutzes? [Are you in favour of extending landscape protection?]	*Wenn es darum geht erneuerbare Energien zu fördern, ist sogar eine Lockerung angebracht.* [When it comes to promoting renewable energy, even a relaxation is appropriate.]	AGAINST	0.006
La Suisse devrait-elle engager des négociations pour un accord de libre échange avec les Etats-Unis? [Should Switzerland start negotiations with the USA on a free trade agreement?]	*Il faut cependant en parallèle veiller à ce que la Suisse ne soit pas mise de côté par les Etats-Unis !* [At the same time it must be ensured that Switzerland is not sidelined by the United States!]	AGAINST	0.010

Figure 10: Some classification errors where the predicted probability of the ground-truth label is extremely low

Future Work

- Design effective mechanisms to solve the challenging scenarios where stances are expressed in an implicit or sarcastic way
- Conduct experiments on more cross-lingual stance detection datasets, including sardistance[7] and ans[8]

[7] Cignarella, A. T.; Lai, M.; Bosco, C.; Patti, V.; Paolo, R.; et al. 2020. SardiStance@ EVALITA2020: Overview of the Task on Stance Detection in Italian Tweets. In EVALITA, 1–10.
[8] Khouja, J. 2020. Stance Prediction and Claim Verification: An Arabic Perspective. In FEVER, 8–17. Online.

Thank you !