Aligning Multilingual Embeddings for Improved Code-switched Natural Language Understanding

Barah Fazili, Preethi Jyothi

Indian Institute of Technology Bombay, India

Motivation

- Multilingual pretrained models, while effective on monolingual data, do not transfer well to code-switched text.
- Does cross-lingual alignment help models understand code-switching better?

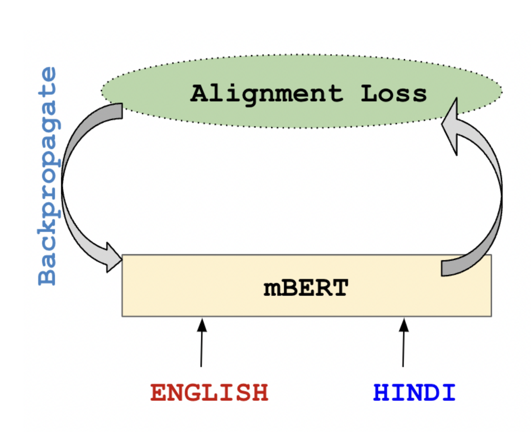
Approach

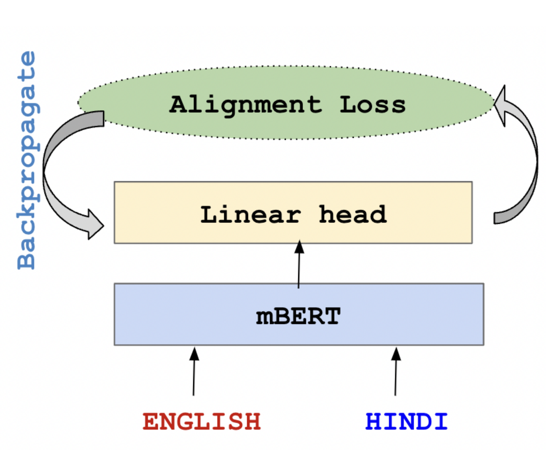
We train multilingual models with alignment objectives using parallel text so as to explicitly align word representations with the same underlying semantics across languages and improve code-switched NLP.

- Parallel bilingual corpus in English and Hindi is used for improving the cross-lingual alignment of multilingual BERT model over the two languages.
- We explore two alignment strategies and achieve improvements of up to 1.16%, 0.76% and 1.9% on code-switched Hindi-English Sentiment Analysis, Named Entity Recognition and Question Answering tasks compared to the baseline mBERT.

Model Design

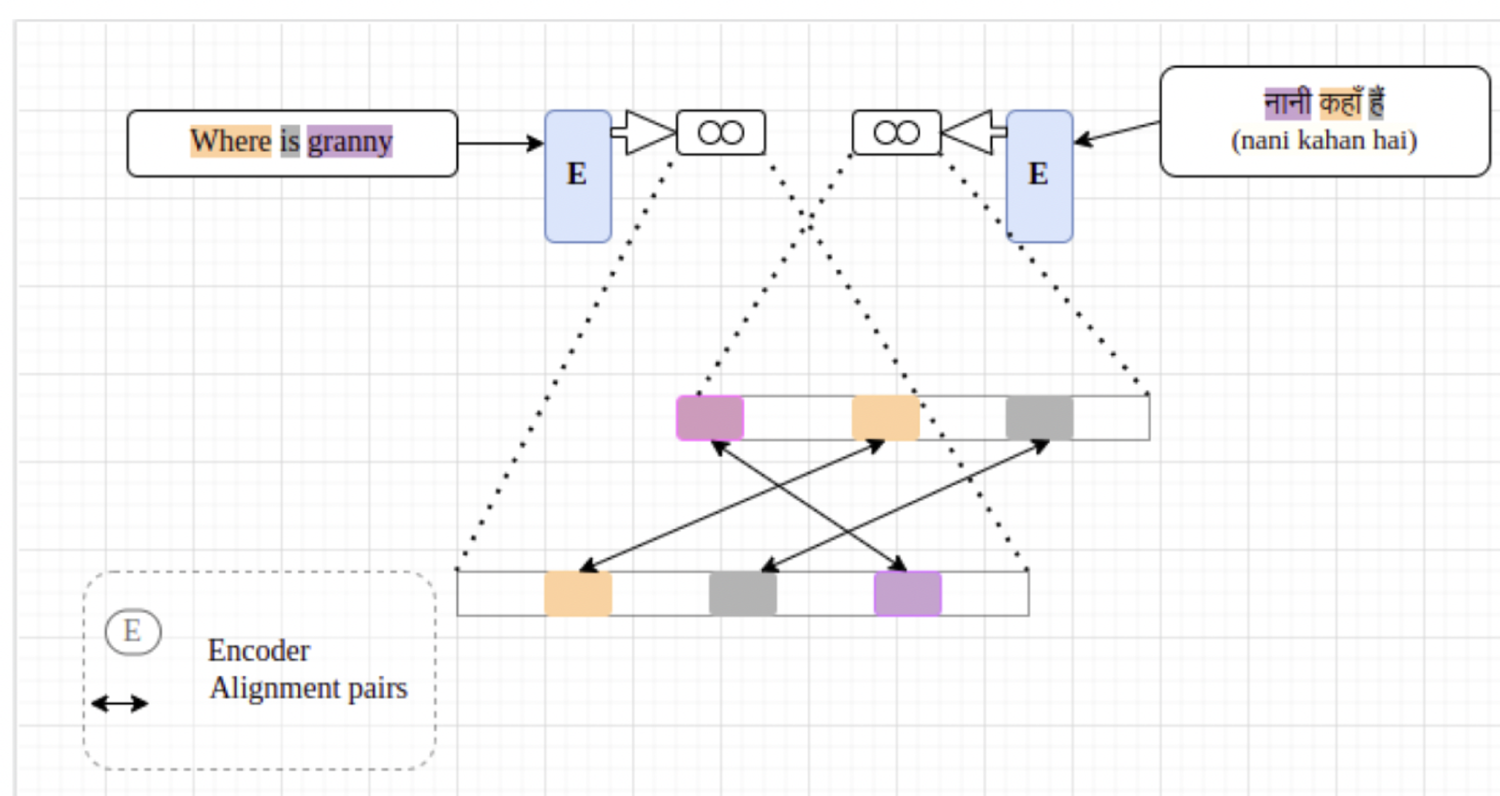
1. Pretrained mBERT which is finetuned to optimize the cross-lingual alignment over the parallel corpus.
2. A linear layer is trained on top of frozen pretrained mBERT.





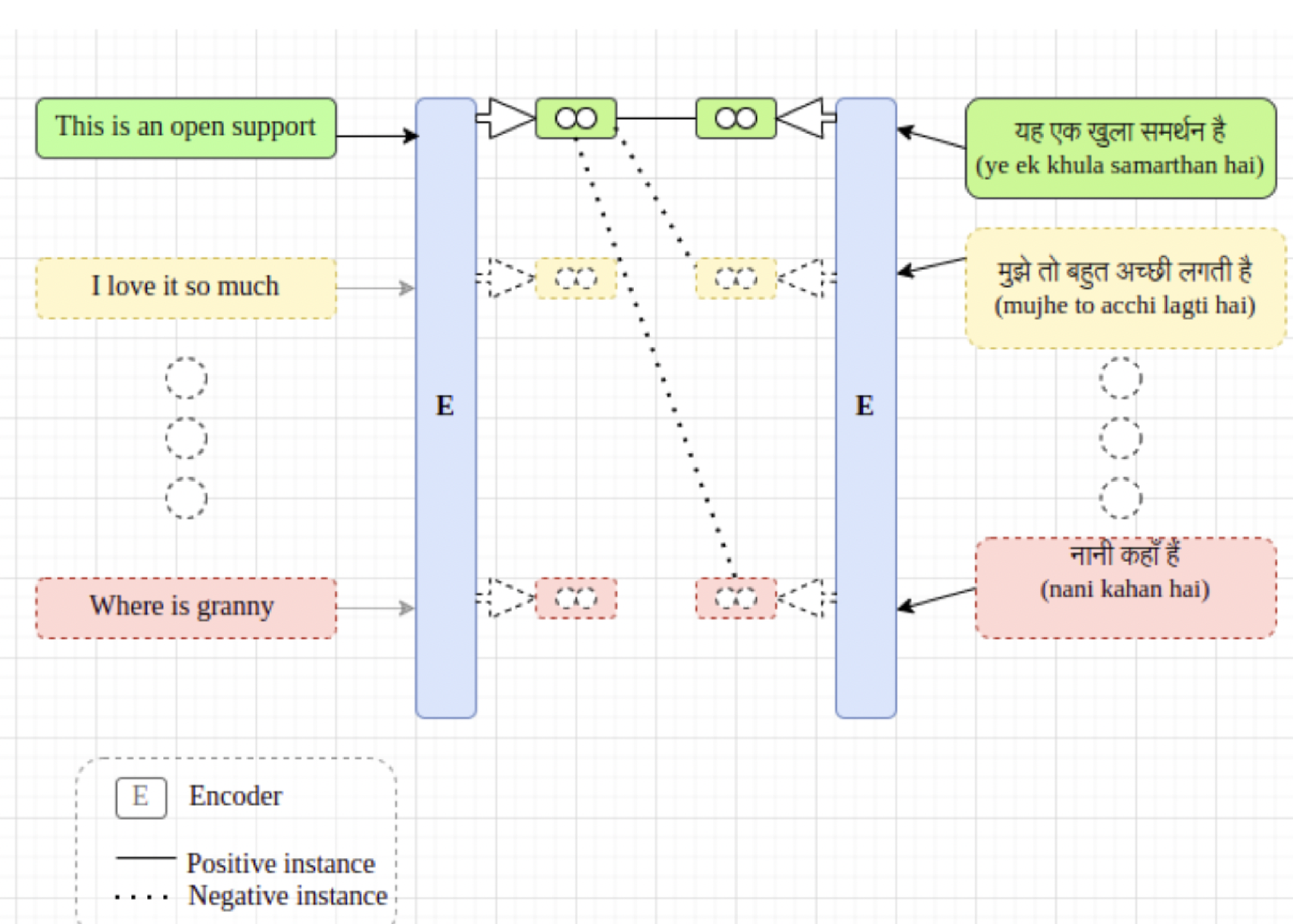
Word-level multilingual alignment

- Word-level alignments across the parallel sentences (in English and Hindi) are first derived using an alignment tool (e.g. awesome-align[1])
- The model is finetuned to further reduce the mean squared distances between contextual embeddings of each such word pair.



Sequence-level contrastive alignment

- Alternatively, the model is optimized using a contrastive objective over a batch of multiple pairs of parallel sentences in En and Hi.
- The loss is computed by summing over the negated cosine similarity between the [CLS] token embeddings of each pair of parallel sentences in En and Hi.



Downstream evaluation of aligned embeddings vs the baseline pretrained mBERT

GLUECoS [2] offers a benchmark on NLP tasks in code-switched Hi-En. Below are average F-scores (± variance) for Sentiment Analysis, NER and Question Answering in code-switched Hindi-English.

SA_EN_HI	
Baseline	60.3 ± 0.00
Contrastive loss(||Devanagari)	59.4 ± 0.01
Multilingual loss(||Devanagari)	**61.0 ± 0.01**
Contrastive loss(MUSE Devanagari)	60.7 ± 0.0
Multilingual loss(MUSE Devanagari)	59.6 ± 0.01

System	NER	QA
Baseline	78.7 ± 0.01	73.5 ± 2.75
Baseline (random)	78.5 ± 0.01	71.8 ± 1.67
Contrastive loss(||Roman)	**79.3 ± 0.00**	74.0 ± 2.2
Multilingual loss(||Roman)	79.0 ± 0.01	74.3 ± 2.99
Contrastive loss(||Devanagari)	79.2 ± 0.00	72.3 ± 0.69
Multilingual loss(||Devanagari)	78.8 ± 0.01	74.0 ± 1.36
Contrastive loss(MUSE Roman)	-	**74.9 ± 1.44**
Multilingual loss(MUSE Roman)	-	72.6 ± 1.91

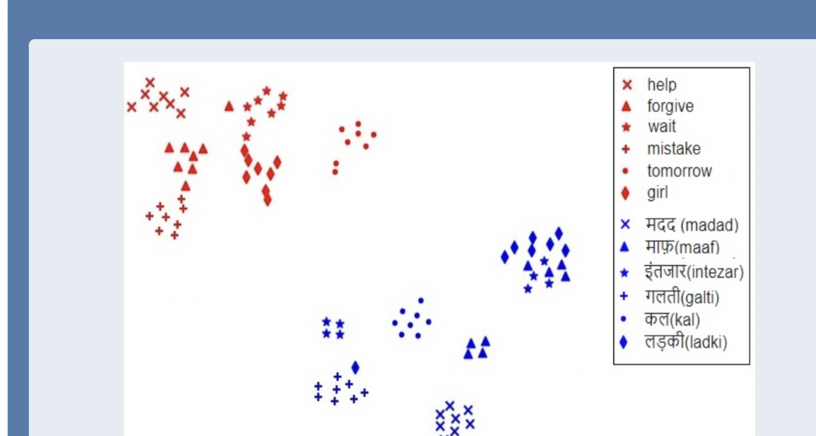
Conclusion

- We propose aligning multilingual embeddings using sentence-level (contrastive) and word-level (non-contrastive) objectives.
- Such an explicit alignment leads to improved performance on three code-switched Hindi-English NLP tasks: SA, NER and QA.
- This method only uses the monolingual resources (parallel text or bilingual lexicon) without demanding pretraining code-switched data.

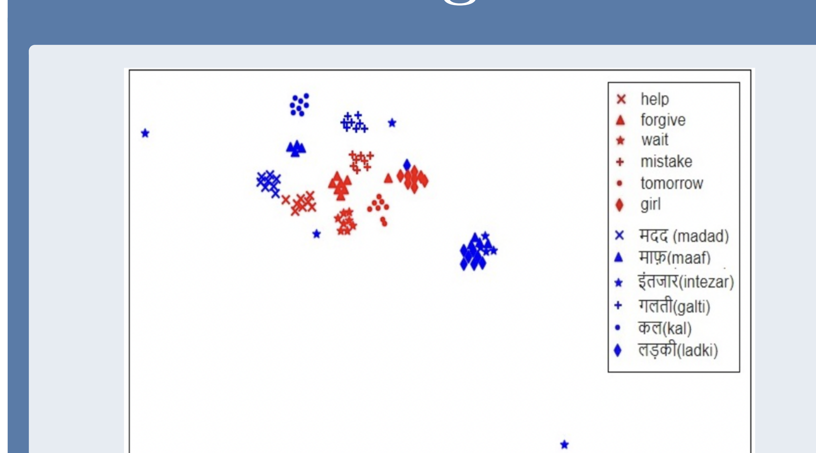
Visualizing the contextual embeddings

- Consider 6 pairs of parallel Hindi and English words appearing in 8 parallel sentences in Hindi and English, respectively.
- The contextual embeddings from mBERT after alignment show the parallel words coming closer together.

Pre-alignment



Post-alignment



References

[1] Zi-Yi Dou and Graham Neubig. Word alignment by fine-tuning embeddings on parallel corpora, 2021.

[2] Simran Khanuja, Sandipan Dandapat, Anirudh Srinivasan, Sunayana Sitaram, and Monojit Choudhury. GLUECoS : An evaluation benchmark for code-switched NLP, 2020.