


Motivation

Building dialogue systems that cater to different languages is of utmost importance in both academic and commercial settings.

But, dialogue data collection for multiple languages is expensive, tedious, and requires native supervision.

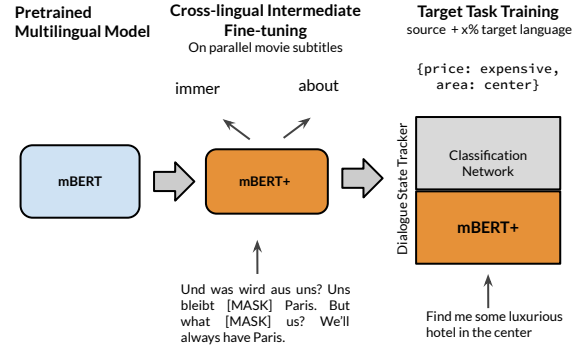
Instead, "transfer" the knowledge from one language to another using shared multilingual language models.

But, multilingual language models are trained with data which is different than human conversation.

 **Fine-tune the multilingual language model using parallel and conversational data with cross-lingual intermediate tasks and then fine-tune for downstream dialogue tasks.**

Our cross-lingual tasks are based on interaction between:
(i) source and target language
(ii) dialogue history and current response

Proposed Method



Cross-lingual Intermediate Tasks

Task: Predict the [MASK] tokens

English subtitle

Who is it, Martin? A bat, Professor. Don't waste your pellets. You'll never harm the bat.

German subtitle

Wer ist denn da, Martin? Eine Fledermaus, Herr Professor. Verschwinden Sie kein Schrot darauf. **Dieser Fledermaus können Sie nichts anhaben**

XDM

Who is it, Martin? A [MASK] Professor. Don't waste your pellets **Dieser Fledermaus [MASK] Sie nichts anhaben**

Response Masking

Who is it, Martin? A bat, Professor. Don't waste your pellets **[MASK] [MASK] [MASK] [MASK] [MASK] [MASK]**

TLM

Who is it, Martin? A [MASK], Professor. Don't waste your pellets. You'll never [MASK] the bat. Wer ist denn da, Martin? Eine Fledermaus, Herr Professor. Verschwinden Sie kein [MASK] darauf. **Dieser Fledermaus können Sie [MASK] anhaben**

XDM: Cross-lingual Dialogue Modelling TLM: Translation Language Modelling

Experimental Setup

Dataset	Source	Target
MultiWoZ (Gunasekara et al., 2020)	En + 10% Zh	Zh
MultiWoZ	Zh + 10% En	En
MultilingualWoZ (Mrkšić et al., 2017)	En only	De, It

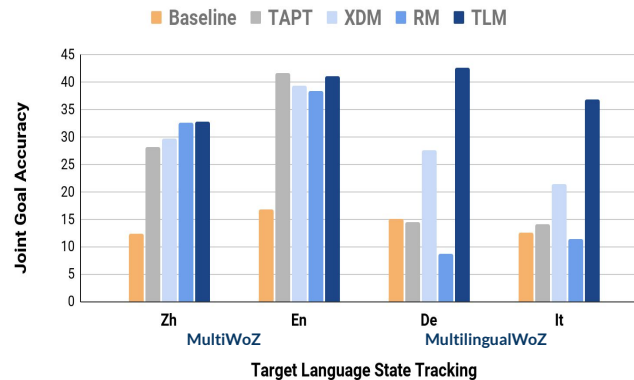
Baseline: mBERT based dialogue state tracker without any intermediate fine-tuning.

TAPT: Task adaptive pretraining - intermediate fine-tuning with chats from the source language dialogue state tracking dataset.

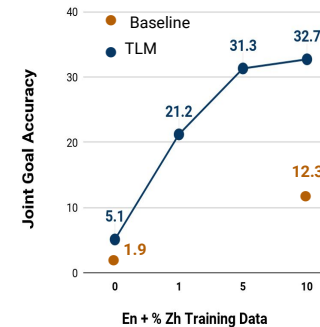
En: English, Zh: Chinese, De: German, It: Italian

Results

Cross-lingual intermediate fine-tuning is helpful



Target Data Efficiency



Insights

TLM based architecture with dynamic code mixed augmentation (Qin et al. 2020) leads to state-of-the-art performance for Multilingual WoZ.

Using chat-style intermediate data (movie subtitles) is better than out of domain intermediate data (news).

Using dialogue history based intermediate tasks is better than using utterance level intermediate tasks.

References

- Overview of the Ninth Dialog System Technology Challenge: DSTC9, Gunasekara et al., arXiv 2020
- Semantic specialization of distributional word vector spaces using monolingual and cross-lingual constraints. Mrkšić et al., ACL 2017
- CoSDA-ML: Multi-Lingual Code-Switching Data Augmentation for Zero-Shot Cross-Lingual NLP, Liu et al., IJCAI 2020

This work was supported by the UKRI Centre for Doctoral Training in Natural Language Processing, funded by the UKRI (grant EP/S022481/1), the University of Edinburgh, and Huawei