Pre-training and Data Augmentation for Dialogue State Tracking

Paridhi Maheshwari, Sharmila Reddy Nangi
{paridhi, srnangi}@stanford.edu

Introduction

- Dialogue State Tracking (DST) is an important step in task-oriented dialogue systems.
- Slot filling based on the current user utterance and dialogue history.

Sys: Hi, what can I do for you?
User: Please find me a Chinese restaurant.
State: *restaurant-food=chinese*

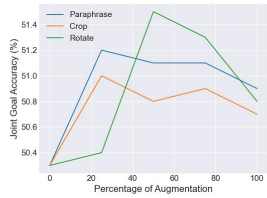
Sys: Charlie Chan fits your criterion, can I book it for you?
User: Yes, I need a table on Monday at 12:15.
State: *restaurant-food=chinese; restaurant-name=charlie chan restaurant-book day=monday; restaurant-book time=12:15*

Sys: Booking is successful. Is there anything else I can assist you with today?
User: I also need a taxi to get me to the restaurant on time.
State: *restaurant-food=chinese; restaurant-name=charlie chan restaurant-book day=monday; restaurant-book time=12:15 taxi-destination=charlie chan; taxi-arriveby=12:15*

- We explore pre-training strategies and data augmentation methods to better leverage the power of large-scale, pre-trained language models (like BERT and T5) for DST.
- Multi-Domain Wizard-of-Oz (MultiWoZ) Dataset [1]
  - Multi-turn dialog utterances with labeled slot-value pairs.
- **Evaluation:** Joint Goal Accuracy, Slot F1, Slot Accuracy.

Data Augmentation for Dialogue Understanding

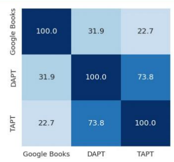
- LMs perform better when provided with huge amounts of training data [3].
- **Why Data Augmentation?**
  - Additional training data without any human annotation effort.
  - Provides diversity for better generalizability of NLP models.
  - Noisy annotations in MutiWoZ dataset (multiple versions).
- **Rule-Based Augmentations**: Entity Replacement, Crop and Rotate using dependency parse trees, Sequential augmentation to increase complexity.
- **Deep Learning Techniques**: Paraphrasing using Pegasus, Reverse translation with English-Spanish NMT.
- Data augmentation significantly improves performance on DST.
- High levels of augmentation can at times hurt the performance too.



Original Utterance	Can you find an Indian restaurant for me that is also in the town centre ?
Entity Replacement	Can you find an Mexican restaurant for me that is also in the town east ?
Paraphrase	I want to go to an Indian restaurant in the centre of the town.
Translate	I am looking for an Indian restaurant that is also in the city center ?
Sequential	Can you find an Indian restaurant for me that is also in the town centre ? I want to make a reservation for two people.
Crop	find an Indian restaurant that is in the town centre
Rotate	an Indian restaurant for me that is also in the town centre find you

Multi-phase Adaptive Pre-training

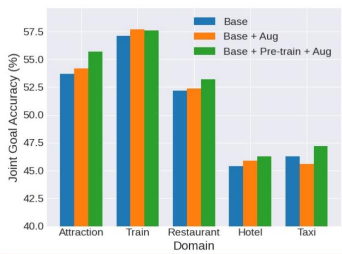
- Significant domain mismatch between large-scale text corpus (for language modelling) and dialogue datasets.
  - Domain adaptive pre-training on open-domain dialogues.
  - Task adaptive pre-training on target dataset.
- Span-level objectives to reason across multiple turns and span-selection where entities are contiguous sequences of words.



Model	Objective	Pre-training	MultiWOZ 2.0 JGA	MultiWOZ 2.0 S F1	MultiWOZ 2.0 SA	MultiWOZ 2.1 JGA	MultiWOZ 2.1 S F1	MultiWOZ 2.1 SA
BERT	Span Prediction	Base	51.2	**75.2**	**92.6**	53.5	75.8	92.2
		+ DAPT	**52.1**	75.1	**92.6**	54.6	**76.3**	**92.3**
		+ TAPT	50.7	**75.2**	92.5	**54.9**	75.9	92.2
		+ DAPT + TAPT	51.7	75.1	**92.6**	53.6	75.9	92.2
T5	Span Corruption	Base	50.2	89.6	96.5	50.3	90.2	96.3
		+ DAPT	50.2	89.4	96.4	50.8	90.2	96.5
		+ TAPT	**51.6**	**89.9**	**96.6**	**51.3**	**90.4**	**96.7**
		+ DAPT + TAPT	51.5	89.6	96.5	50.2	90.2	96.4

Domain Analysis

- Combining best pre-training method and data augmentation techniques.
- Consistent gains in goal accuracy across all domains in MutiWoZ.



Conclusion & Future Work

- Incorporating language structure of dialogues through span-level pre-training and additional domain data through augmentation methods is helpful for DST.
- Potential future direction would be to study the impact of these techniques on end-to-end dialogue systems including generation.

References

[1] Budzianowski et al., Multiwoz-a large-scale multi-domain wizard-of-oz dataset for task-oriented dialogue modelling, EMNLP 2018
[2] Gururangan et al., Don't stop pretraining: Adapt language models to domains and tasks, ACL 2020
[3] Feng et al., A Survey of Data Augmentation Approaches for NLP, ACL 2021