Few shot QA with DNN

Adil Sadik
asadik@stanford.edu

Problem

- Standard QA models outperforms humans but suffers at few shot QA tasks.
- Models tend to overfit to training dataset.
- Labeled data is not abundant and expensive. No or very few labeled data in some domains.
- Need a better QA model to tackle this few shot learning task.

Background

- Key idea is to learn Domain Invariant features.
- Approaches: GAN, Meta learning, TAPT
- Recent work also explored clever MLM objectives, bi-encoder model, encoder-decoder models to learn better contextualized representation of QA task.

Methods

- Pre-train QA model with a new special **[QUESTION] token**. Key idea is to do a better job learning QA task specific features using a new task specific token.
- Use **reptile** [1] ML algorithm to find an ideal initialization point of the model parameters.
- **Modify** reptile algorithm to sample mini-batches from all domains in the inner loop.

Experiment & Analysis
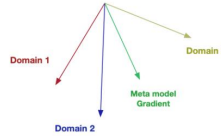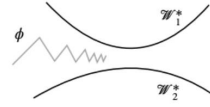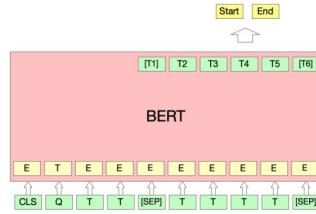
Model	EM	F1
Baseline	32.8	48.8
QA-MLM[1]	30.1	44.1
QA-MLM[2]	35.1	49.9
QA-MLM[3]	32.9	48.6
TAPT-MLM[4]	33	49
PT-META - val	**37.17**	**51.91**
PT-META - test *	40.48	57.71

Val- EM: 13%
Val-F1: 6%
Improvement from baseline

In-domain VAL:
F1 : 70
EM: 55

* Not from the most optimized model.

Conclusions

- Pre-training by adding new vocabulary is effective. It guides model to learn task specific features.
- Meta learning is effective in finding an optimal set of parameters for domain adaptation.

References

- [1] OpenAI reptile: https://openai.com/blog/reptile/