Unsupervised Domain Adaptation for Large Language Models

Essential ideas

Neural Unsupervised Domain Adaptation in NLP - A Survey (Ramponi & Plank, 2020)

We are interested in **Unsupervised Domain Adaptation** for LLMs

- Problem: *source* and *target* distribution do not match:
    - Sometimes called *dataset shift*. In NLP, **domain shift**
- Not *transfer learning*, but can be seen as a particular case of it:
    - *Transductive transfer learning* (see for example *Sebastian Ruder's Thesis - Neural Transfer Learning for Natural Language Processing, 2019*)
- We define a **domain** $\mathcal{D} = \{\mathcal{X} | P(X)\}$ and a **task** $\mathcal{T} = \{\mathcal{Y}, P(Y|X)\}$
    - Assume that *source* and *target* tasks $\mathcal{T}_S$ and $\mathcal{T}_T$ stay the same
    - Assume a shift in *domain distributions* $\mathcal{D}_S$ and $\mathcal{D}_T$ - but we make **no assumption on labels** $\rightarrow$ Covariate shift !

Essential ideas

Neural Unsupervised Domain Adaptation in NLP - A Survey (Ramponi & Plank, 2020)

In NLP, what is a domain ?

- Some sort of coherent type of corpus ?
  - We know of non-canonical kinds of data that need specific models (TweetBERT, SciBERT, BioBERT...)
  - What about languages ?
  - Frequency, social variations, data collection strategies ...
- Idea of *variety space (What to do about non-standard (or non-canonical) language in NLP, Plank, 2016):* A domain forms a region in that space, with some dimensions being various aspects of language variation
- Related problems:
  - Cross-lingual learning: extreme DA with parallel data
  - Domain generalization: DA to several domains
  - Out-of-domain generalization: robust everywhere !

Today

Neural Unsupervised Domain Adaptation in NLP - A Survey (Ramponi & Plank, 2020)

Hence, we look into *unsupervised* and *task-independant* approaches

Division into **model-centric** and **data-centric** approaches

- Model-centric: re-design partof the model - feature, loss
- Data-centric: pseudo-labeling, data selection .... but mainly,
- **Pre-training approaches:** with adaptive secondary steps,
  - Adaptive pre-training: continuing with domain specific (multi-phase) pre-training

    (Unsupervised Domain Adaptation of Contextualized Embeddings for Sequence Labeling, Han and Eisenstein, 2019) + (Don't Stop Pretraining: Adapt Language Models to Domains and Tasks, Gururangan et al, 2020)
  - Auxiliary task pre-training: usually supplementary training on intermediate labeled-data task

    (Rethinking Why Intermediate-Task Fine-Tuning Works, Chang et al, 2021)

Today

Promising direction for domain adaptation: using a dedicated, lightweight layer; the most famous approach being the **adapter**

(Parameter-Efficient Transfer Learning for NLP, Houlsby et al, 2019)

- Already used for cross-lingual learning *(MAD-X: An Adapter-Based Framework for Multi-Task Cross-Lingual Transfer, Pfeiffer et al, 2020)* and adding knowledge to models *(K-ADAPTER: Infusing Knowledge into Pre-Trained Models with Adapters, Wang et al, 2021)*

Our recent work on **SNCF conversational data** uses a mix of most of these ideas !

An Adaptive Layer to Leverage both Domain and Task Specific Information from Scarce Data (Guibon et al, 2023)