

Zero-shot, one-shot, few-shot

PA154 Language Modeling (12.1)

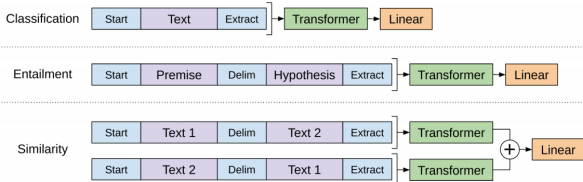
Pavel Rychlý

pary@fi.muni.cz

May 4, 2023

Usage of Large Models

- training of big models on huge data is expensive (long training time)
- fine tuning on small data of target task
- combining language model with additional NN/layer, training only new layer
 - big model is frozen, only used



Usage of LLM without training

Usage of LLM without fine tuning

- fine tuning is still expensive
- models can predict reliably
- using generation for end tasks
- zero-shot no task-specific data/training

Zero-shot

- formulate task in natural language
- task description + prompt
- Translate English to French:
cheese =>
- Summarize the following paragraph into one sentence:
text

One-shot

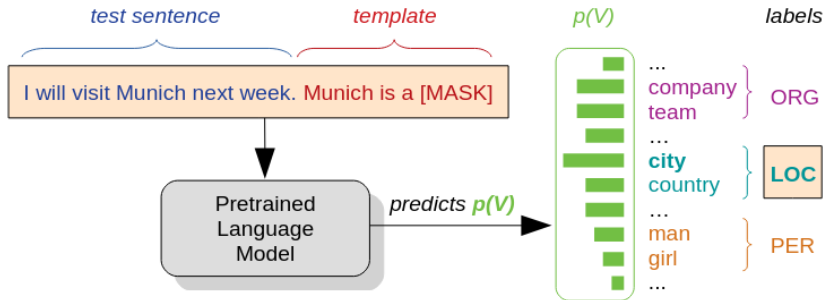
- formulate task in natural language and show one example
- task description + example + prompt
- Translate English to French:
seq otter => loutre de mer
cheese =>

Few-shot

- formulate task in natural language and show a few examples
- task description + examples + prompt
- Translate English to French:
seq otter => loutre de mer
peppermint => menthe poivrée
plush girafe => girafe peluche
cheese =>

Language Models are Few-Shot Learners

Token is MASK



TOKEN is a MASK: Few-shot Named Entity Recognition with Pre-trained Language Models

Available LLM

- Pythia
- OPT
- GALACTICA
- T0