Large Language Models - Seminar

Topics

Marion Di Marco

Cross-Lingual Abilities of PLMs

- Multilingual LMs (e.g. mBERT) learn rich cross-lingual representations
- How much word-level translation information is embedded in mBERT and how can it be extracted?
  - template-based querying for word translations
  - analogy-based translation
- How does mBERT represent language information?
  - hypothesis: mBERT representations contain a language-encoding component and a language-neutral component
  - how can these representations be disentangled?

Cross-Lingual Abilities of PLMs

It's not Greek to mBERT:
Inducing Word-Level Translations from Multilingual BERT

Hila Gonen, Shauli Ravfogel, Yanai Elazar, Yoav Goldberg (2020)

Proceedings of the Third BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP

`https://aclanthology.org/2020.blackboxnlp-1.5.pdf`

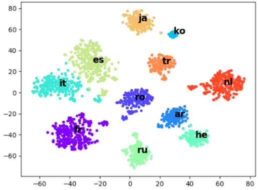


Figure 1: t-SNE projections of the representations of the template-based method.

Linguistic Structure in LLMs

- LLMs perform very well at many language tasks
- To what extent can these abilities be attributed to generalizable linguistic understanding vs. surface-level lexical patterns?
- Apply structured prompting in autoregressive language models
- Look at word-level and span-level sequence tagging tasks
  - POS tagging, sentence chunking, Named Entity Recognition
  - zero- and few-shot settings

Prompting Language Models for Linguistic Structure

Terra Blevins, Hila Gonen, Luke Zettlemoyer (2023)

Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics

https://aclanthology.org/2023.acl-long.367.pdf

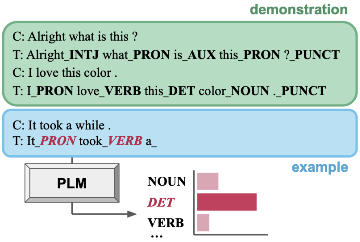


Figure 1: Sequence tagging via structured prompting. Each predicted label is appended to the context along with the next word to iteratively tag the full sentence.

Subword Segmentation

- LLMs operate on subword tokens to handle OOVs and for efficiency
- Subword segmentation strategies (BPE, WordPiece, ...): compression algorithms based on (sub) word frequencies, do not take into account linguistic information
- Example segmentations from BERT

  `realize` → `realize`
  `finalize` → `final, ##ize`
  `mobilize` → `mob, ##ili, ##ze`

- Linguistically inconsistent, subwords are often meaningless
- What does this mean for BERT's ability to generalize and to infer the meaning of complex words?

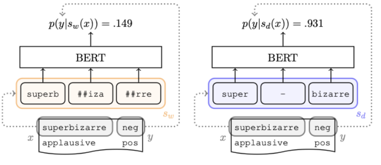
Subword Segmentation

Superbizarre Is Not Superb: Derivational Morphology Improves BERT's Interpretation of Complex Words

Valentin Hofmann, Janet Pierrehumbert, Hinrich Schütze

Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing

https://aclanthology.org/2021.acl-long.279.pdf



(a) BERT (s_w) (b) DelBERT (s_d)

Figure 1: Basic experimental setup. BERT with WordPiece segmentation (s_w) mixes part of the stem `bizarre` with the prefix `super`, creating an association with `superb` (left panel). DelBERT with derivational segmentation (s_d), on the other hand, separates prefix and stem by a hyphen (right panel). The two likelihoods are averaged across 20 models trained with different random seeds. The average likelihood of the true class is considerably higher with DelBERT than with BERT. While `superbizarre` has negative sentiment, `applausive` is an example of a complex word with positive sentiment.