



Augmenting Language Models for Commonsense Generation

Modern deep neural networks consist of huge number of parameters. Recent research shows that the parameters are mostly used to “store” external knowledge. For example, GPT-3, which consists of 1,750 billion parameters, has surprisingly good performance in downstream applications, especially knowledge-extensive tasks including commonsense generation, etc.



Recently, researchers propose to fetch knowledge from external knowledge bases, instead of storing all the knowledge in the model parameters. Such method generalizes better as it is able to retrieve knowledge from any knowledge bases, including Wikipedia, and even the entire web.

In this thesis, we investigate the task of commonsense generation by leveraging the retrieval augmentation technique. You will use state-of-the-art technologies for information retrieval and text generation to build language models for generating commonsense inferences.

Requirements: Strong motivation, knowledge in deep learning, or a solid background in machine learning. Experience with Python and PyTorch is an advantage.

Interested? Please contact us for more details!

Contact

- Zhao Meng: zhmeng@ethz.ch, ETZ G60.1