

Computational Cognitive Modeling - Spring 2023

DS-GA 1016 / PSYCH-GA 3405.004 Project Proposal

Harsh Dubey (hd2225@nyu.edu)

Rahul Meghwal (rm5707@nyu.edu)

Yulu Qin (yq810@nyu.edu)

Examining Lexical and Syntactic Structure in BERT/GPT

With the release of GPT-4, a transformer-based large language model, many are struck by its ability to generate correct sentences and understand complex ideas. However, as with other transformer-based language models, there are concerns about its potential risks since it is still opaque, despite its impressive functionality.

Meanwhile, the process of how humans learn language and form lexical and syntactic structures remains a mystery. Some researchers suggest that such rapid progress in NLP has the potential to transform debates about how humans learn language (Bowman, 2022).

Elman's seminal work in 1990 showed how Simple Recurrent Networks can learn meaningful syntactic and semantic representations without targeted inductive biases. Since then, the NLP community has continued this line of research. Linzen found that long short-term memory (LSTM) language models are able to capture subject-verb agreement in many common cases (Linzen, 2016). Rogers et al. trained BERT on larger-scale written text corpora and examined its linguistic representations.

In this project, we aim to further explore the underlying semantic and syntactic representations that state-of-the-art language models (such as BERT and GPT-3) may incorporate. Inspired by Elman's hierarchical clustering analysis, we want to examine the hierarchical nature of the learned representations after fine-tuning the models on a domain-specific dataset. In terms of syntax, we will follow the subject-verb agreement task to examine the models' syntactic understanding and make comparisons.

Since the pre-trained GPT-4 model is not available to the public for fine-tuning and testing, we may conduct linguistic analyses by simply interacting with it through online communications and testing its linguistic understanding under certain tasks.