

Assignment 5

CMSC 473/673 — Introduction to Natural Language Processing

Due Wednesday December 15th, 11:59 PM

Item	Summary
Assigned	Monday December 6th
Due	Wednesday December 15th, no late days after December 20th
Topic	Transformers and Linguistic Phenomena
Points	50

In this assignment you will use pre-trained large language models to explore how well they capture different types of syntactic and/or semantic phenomena. There is no training of these models required.

You are to *complete* this assignment on your own: that is, the code and writeup you submit must be entirely your own. However, you may discuss the assignment at a high level with other students or on the discussion board. Note at the top of your assignment who you discussed this with or what resources you used (beyond course staff, any course materials, or public Discord discussions).

The following table gives the overall point breakdown for this assignment.

Question	1
Points	50

What To Turn In Turn in a writeup in PDF format that answer the questions; turn in all requested code necessary to replicate your results. Be sure to include specific instructions on how to build (compile) your code. Answers to the following questions should be long-form. Provide any necessary analyses and discussion of your results.

How To Submit Submit the assignment on the submission site:

<https://www.csee.umbc.edu/courses/undergraduate/473/f21/submit>.

Select “Assignment 5.”

Full Questions

1. (50 points) In this question, you will put on a “probing” (linguist/science-based) hat. You will think about different types of syntactic and/or semantic phenomena, and see how well large language models (Transformer methods) can capture them, via notions of embedding similarity.

- **Your Task** Identify three different linguistic phenomena: call these P_1, P_2, P_3 . For each phenomena P_i , construct a minimum of $M = 5$ pairs of sentences that address or evince that phenomena. Use a pretrained, large language model (sometimes abbreviated LLM) to evaluate the similarity of each pair. Create and turn in a report in which you:
 - Provide the M sentences for each phenomena.
 - Discuss how you are using evaluating the similarity of each pair.
 - For each phenomena, present your similarity results. This could be done in a graphical form (bar or scatter is probably the most understandable among graphical options), tabular form, or both. (You don’t have to do both, but if you feel that doing so helps you present your results, then you can.)
 - For each phenomna, summarize and analyze the results. Your analysis should discuss, based on this limited exploration, your take-aways on how well LLMs can capture these phenomena, and where they fall short.

Turn in all code, including what is necessary for producing the plots.

- **How You Will Be Graded** Your grade on this problem will be on how well your report completes the above tasks.
- **Specifics** You may use the extended Colab notebook (https://colab.research.google.com/drive/1nE5rpk1kdz3GBDzjq0Pg4k_7UuxmfZV7?usp=sharing) we had in class as starter/inspiration.

The choice of which LLM to use is up to you, with one exception: you may **not** use either `bert-base-uncased` or `bert-base-cased`.

The choice of how you evaluate similarity it also up to you.

- **Example** Let’s say you were examining the following “a” vs. “an” phenomena:

“a” vs. “an”: Generally, we use “a” before nominals that do not begin with a vowel (sound) and “an” before nominals that do. For example, we say “a banana,” and “an apple,” but we also say “a red apple.”

For this phenomena, you’d want to construct at least 5 pairs of phrases that focus on this syntactic rule. You want each pair to be appropriately contrastive. This contrast could be between (grammatically) acceptable and not [for example, “an apple” vs. “a apple”], or its interaction with other grammatical rules [e.g., “an apple” vs. “a red apple,” or “an apple” vs. “an inedible apple”], or its interaction with spelling vs. sound [e.g., “an hour” vs. “a hour”], or other interactions *that you think of*. For each phrase, use a LLM to embed that phrase and compute how (dis)similar the LLM judges those phrases.

I would *highly* recommend focusing one of these interactions and coming up your 5 pairs based on that single interaction. That is, don’t try to evaluate grammatical acceptability and grammatical rule interaction *and* spelling vs. sound. Pick one and focus on that.

- **Possible Phenomena** The following are *some* possible phenomena. **You are not restricted to this list**, with the exception that you may not use the “a vs. an” phenomena (since it was given as an example). You may also **not use any examples given here**.

Adjective Order There are many different ways that adjectives can modify nominals. Some modify the number (“two”), others modify a subjective value (“great”), and other modify physical/observable attributes (e.g., size, color and shape). With these modifiers, it is more natural to say “two great big green houses” than it is to say “great green two big houses.”

Transitive vs. Di-Transitive vs. Intransitive Verbs Certain verbs are transitive, meaning they can take direct objects (“Chris ran the marathon”); others are di-transitive, meaning they can take direct objects and indirect objects (“Chris gave Pat the book”); and others are intransitive, meaning they do not take any objects (“Chris ran toward the hills”). These may take certain modifiers, e.g., “Chris ran toward the hills with vigor,” “Chris ran the marathon with vigor,” and “Chris gave the book to Pat.”

Noun-Verb Agreement One of the most common instance of English’s inflectional morphology is in noun-verb agreement. For example, “she loves NLP” vs. “they love NLP.”

Selection Preferences Certain verbs have preferences for certain types of arguments. For example, “drink” prefers its *subject* to be animate and its *object* to be drinkable, or at least consumable, as seen in the sentence “Chris drank the soda.” Contrast this with the sentence “Chris drank the laptop,” which just does not make sense. However, these aren’t requirements: imagine getting poison ivy and putting lotion on the rash. You may say “the rash really drank the ointment.” Similarly, some combinations just are more acceptable than others: “Chris drank the soda” is more acceptable than “Chris drank the gravy.”

Upward vs. Downward Entailment This concerns the inferences that can be made by using different types of quantifiers and determiners in sentences, and how they interact with different phrases in the sentence. For example, when we say “Some children like vegetables,” and we focus on the interaction between the quantifier “some” and the noun “children,” we can infer that “A child likes vegetables,” but we cannot infer that “All children like vegetables.” On the other hand, if we say “No child likes steamed vegetables,” and we focus on the interaction of the determiner “no” and the phrase “steamed vegetables,” we can infer that “no child likes steamed broccoli,” but we cannot infer that “no child likes vegetables.”

Veridicality/factuality This concerns how completion of an event, or lack thereof, can be communicated. For example, “Pat ran a marathon” means that Pat actually ran the marathon (and completed it), while “Pat tried to run a marathon” does not necessarily mean that Pat actually completed it.

Deverbal Events Many times we communicate an event via a verb, for example “Pat and Chris passionately argued over the plan.” However, we can also communicate (similar, but not necessarily the same) information about an event through a nominalized form of that verb, such as “Pat and Chris’s argument over the plan was passionate,” or “Pat and Chris’s argument surprised people.”

Subjective Adjectives This concerns how adjectives can interact with one another. In many cases, English adjectives are subjective: if we think of the noun being modified as a set, adjectives identify various subsets of that set, and they act intersectionally. For example, imagine the noun “cats” referring to a set of all possible cats; then “fluffy cats” is a subset of the “cat” set, and “white, fluffy cat” is a further (intersectional) subset. However, not all adjectives, such as “fake,” are subjective. For example, “fake cats” would not refer to a subset of the “cats” set.