

## CS288 HW4 Report

Even though we fine-tuned GPT-2 medium model on the `ted_talks_iwslt` in order to improve its performance in translation from Dutch to English, it still may not do well. Some reasons why I think it didn't do well are the following.

The `ted_talks_iwslt` dataset includes language from TED talks, which may have specific domain-specific terminology and expressions that the model is being fine tuned on. So, if the sentence we're trying to translate does not contain such language, the model may struggle to produce accurate translations.

The `ted_talks_iwslt` dataset is relatively large, but still it may not include examples of every possible sentence structure or vocabulary that the model might encounter. As a result, the model may not be able to accurately translate the types of sentences that were not in the dataset.

Also, even after being finetuned, the GPT-2 medium model is a language model that processes text one word at a time. This means that it may struggle to understand the broader context of a sentence or document, which can lead to errors in translation even if it is finetuned.

### **2 cases where the translation model does not do well is:**

- German: Das Schlossgespenst schwatzte geschwind und geschwätzig in der Schlosskapelle.  
English: The ruins of a sunken city lie under the surface of the ocean.
- German: Ik ben naar de bibliotheek gegaan om een boek te lenen  
English: When you read a book, you're not just reading it

In the first sentence, the model does not do well because the given sentence is a long and complex sentence that is hard to translate with not given a lot of training data and a model that is not domain specific.

In the second sentence, even though the sentence is in the training set, the model does not do well which I think it's because the model wants to do what it was trained on more data to do which is generate fluent sentences. So, even though, it does translate some words in the sentences, it still wants to just generate fluent sentences that are actually unrelated to the input because that's what it was designed to do.

### **2 cases where the translation model does well is:**

- German: Ich habe Zerebralparese.  
English: I have cerebral palsy.
- German: Ik ben blij.  
English: It's am very happy

In the first sentence, the model does really well because this exact sentence is in the training set.

In the second sentence, the model does reasonably well because the sentence is simple and does not include very complex language or words.