

1. Run the implementation of machine translation (MT) with seq2seq and attention at https://pytorch.org/tutorials/intermediate/seq2seq_translation_tutorial.html
 - a. Download data files from <https://download.pytorch.org/tutorial/data.zip> and place it as instructed
 - b. Revise the function filterPairs in the section "Loading data files" as follows:
 - i. filter sentence pairs whose length is less than MAX_LENGTH ~~and which start with eng_prefixes (e.g. "I am", "you are")~~
2. Answer the following questions about the MT model of Question 1
 - a. Which RNN model is used as encoder?
 - b. Which RNN model is used as decoder?
3. Modify the MT implementation to measure the model performance with BLEU:
 - a. Split the sentence pairs as follows:
 - i. Randomly select 10% of the pairs for testing/evaluation
 - ii. For each English sentence in the selected pairs, find all its French translations from the whole set of sentence pairs
 - iii. Make *test data* as the list of (English sentence, list of its French translations)
 - iv. Select all sentence pairs whose English sentences are not included in the test data, as *training data*
 - b. Train the model of Question 1 with the training data
 - c. Evaluate the trained model with the test data by using NLTK library
 - i. https://www.nltk.org/modules/nltk/translate/bleu_score.html
4. The link of Question 1 implements the greedy decoding method. Change it to the beam search decoding method as follows:
 - a. Use <https://github.com/budzianowski/PyTorch-Beam-Search-Decoding>
 - b. Beam size = 10
 - c. You may skip returning decoder attention outputs, as this tutorial would not use them
5. Answer the following questions about the parameters of the beam search decoding method implemented at the <https://github.com/budzianowski/PyTorch-Beam-Search-Decoding>
 - a. What stopping criteria are used?
 - b. What normalization is used?
6. Alternate the above two parameters to achieve higher performance than the greedy decoding method