Axel Dumont
axel.dumont@ensta-paris.fr

1 Question 1

Our greedy decoding method builds a sentence word-by-word. This prevents modeling long-range dependencies but is faster in terms of computation. Beam search explores K candidate sentences in parallel, expanding each by adding all possible next words at each step. It keeps the K most probable candidates. Beam search produces better results but requires more computation as it considers many more options.

2 Question 2

The main issue is that the generated sentences tend to continue indefinitely without properly ending. For example, there may be repeated punctuation like periods and repeated words within a sentence, but the end-of-sentence (EOS) token never appears to properly conclude the sentence.
There might be a few solutions to fix that problem :

- Explicitly force decoding to stop once <EOS> is generated rather than relying only on the max length cutoff.
- Increase the maximum output length limitation to ensure the model has enough steps to generate the <EOS> token.
- Manually insert <EOS> tokens in the decoded output once the repetition begins.

According to paper [6], this is a common problem in NMT known as *lack of covering* inducing *over-translation* A solution might be to implement a coverage vector, sequentially updated during the decoding process to keep track of attention history

3 Question 3

Figures 1 and 2 illustrate the attention coefficients for some of the model's translations. The coefficients are high for corresponding words between the source and target sentences, like 'student' and 'étudiant'. However, the effect of adjective-noun inversion is not well captured, as seen for the phrase 'red car' translated to 'voiture rouge'. The attention weight for 'red' is actually higher for 'voiture' rather than 'rouge', and 'car' has a stronger link to 'rouge' rather than the proper translation. This indicates the model's attention mechanism is not perfectly aligning source and target words in cases of structural differences like inverted adjective-noun order. While it is properly translating individual words, the model seems to struggle with aligning more complex syntactic relationships that require reordering the sentence structure. Further refinement of the attention mechanism could improve its ability to handle these more difficult cases of non-literal word-for-word alignment.

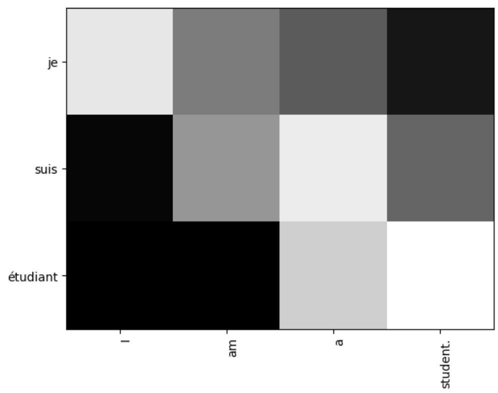

Figure 1: Source/Target alignment for the first sentence.

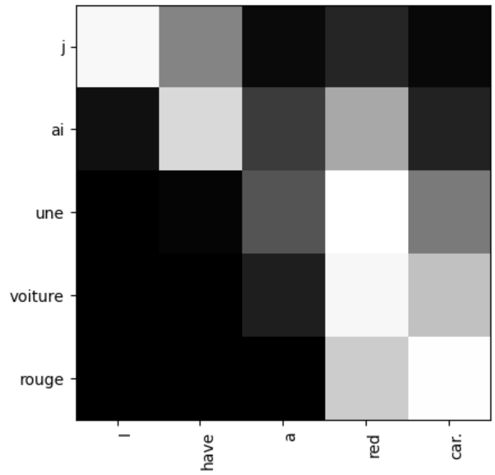

Figure 2: Source/Target alignment for the first sentence.

4 Question 4

An interesting observation is that the word 'meaning' is correctly translated in both sentences, even though it conveys very different meanings in each case. In the first sentence, 'meaning' is a verb meaning 'to intend,' while in the second sentence it is an adjective meaning 'unkind.' This demonstrates the model's capacity to handle homonyms - words that are spelled the same but have different meanings and parts of speech depending on context. The translation of 'meaning' to the appropriate French verb 'vouloir' or adjective 'méchant' indicates the model can disambiguate multiple meanings for a single word based on the surrounding context. Rather than simply translating 'meaning' in isolation, the model understands how to map the same surface form to different underlying meanings. This illustrate the polysemy of the model.

References