

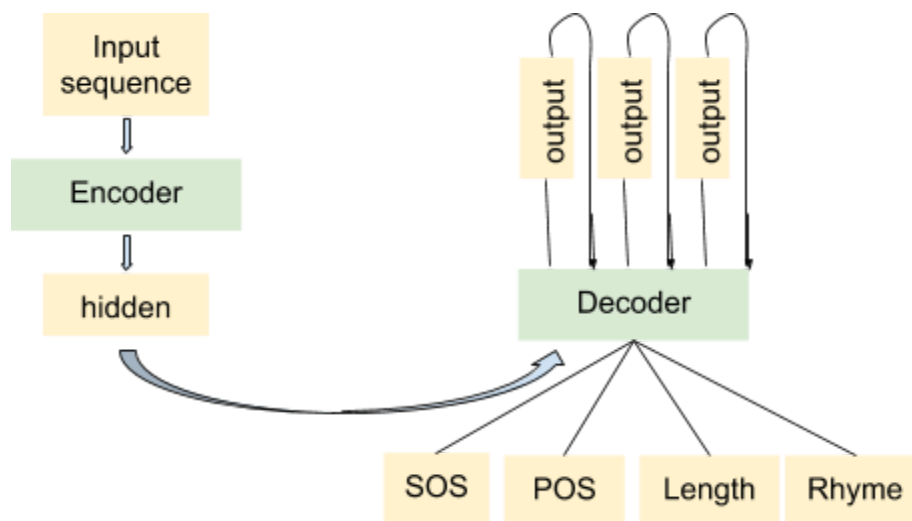
SDML HW3 Task1

B04902105 資工四 戴培倫

Data:

From Task 0, I found that it is really time consuming and hard to train the model using all training set. Therefore, I chose only the first 1000 sentences for target_length between 1 and 31 since only few data have long target.

Models:



- Length
I input 'length left' for every timestep, go through an embedding, concatenate with the input.
- POS
I input 'next POS' for every timestep, go through an embedding, concatenate with the input.
- Rhyme
I input 'target rhyme' at the last timestep and 0 for other timesteps, go through an embedding, concatenate with the input.

The hidden size for sequence is 300, the embedding size for length, pos, rhyme are 10, 10, 20, respectively.

For '**length**' and '**POS**', simply training leads to excellent accuracy as mentioned below.

Length can easily reach almost **100%** accuracy, and POS accuracy can also significantly exceeds the baseline. However, I got poor accuracy on Rhyme, at around 20% accuracy. Therefore, I tried to use method similar to beam search, which is picking the correct rhyme at the last timestep with a certain beam size. At beam size = 50, I can reach 82% accuracy. And by increasing the **beam size to 70**, I can reach **92%** accuracy. *The LM score does not drop after picking the correct rhyme.*

The difference of the two models is simply using the Luong's attention method. I tried the 'dot' method and 'general' method (transforming encoder output). Both methods have similar result, so I chose 'dot' method in order to save resource.

	Length	POS	Rhyme	LM
Attention	0.99954	0.83794	0.92752	31.92072
Non-Attention	0.99970	0.82321	0.92684	30.62061