

Ex2 – Part 2 – Report

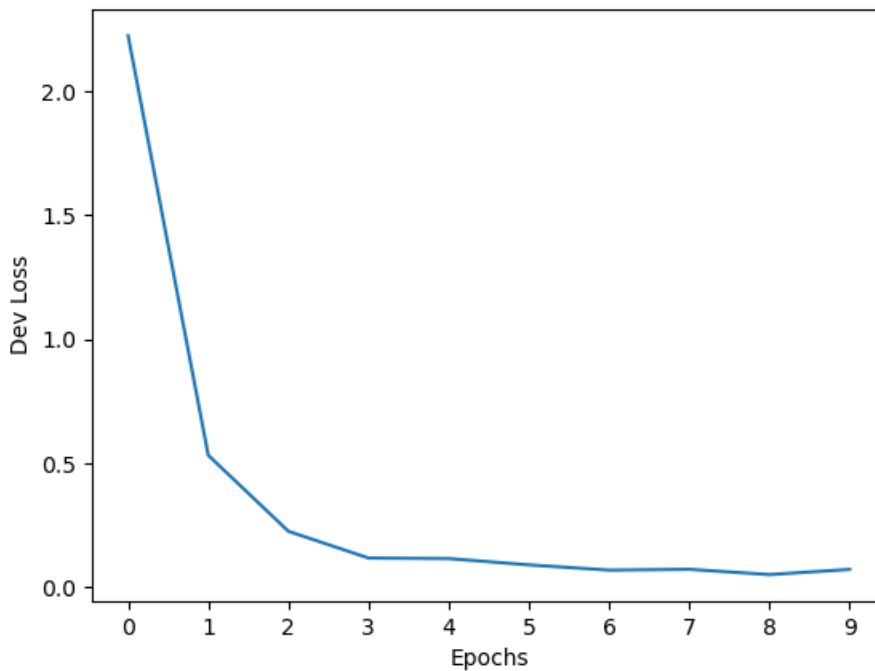
Model's config:

- Criterion: NLL-Loss
- Epochs: 10
- Learning Rate: 0.0003
- Batch Size: 1
- Dropout Rate: 0.3
- Number of LSTM layers: 1
- Embedding Size: 300
- Hidden Size: 512
- Optimizer: Adam
- RNN type: LSTM

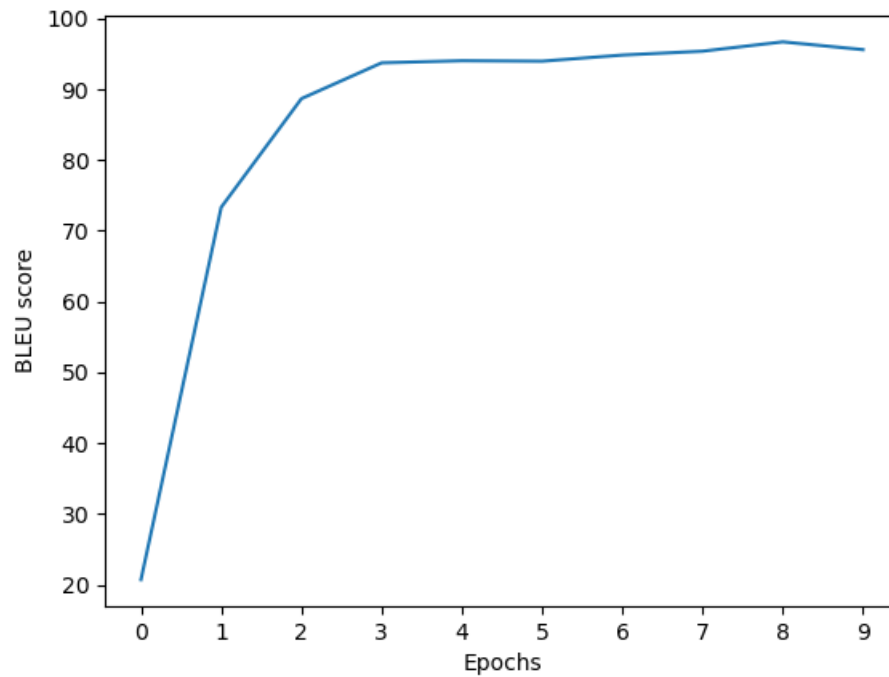
Max sentence length: 15 words

Best Train Loss: 0.0526

Best Dev Loss: 0.0518



Best BLEU score: 96.7



Total runtime (on GPU): 162 seconds (~ 2.7 minutes)

Test BLEU score: 98.718

Test Best Loss: 0.0265

Overall, we can tell that the attention-based model performs a lot better than the Vanilla RNN-based Encoder-Decoder with an improvement of ~14 BLEU scores on the Dev set and ~14 BLEU scores on the Test set.

In manor of run time, the attention-based model took ~10 seconds longer than the RNN-based for the training and ~0.2 seconds longer for the evaluation.