Assignment 4

Your name:
Your GTID:

Seq2Seq Results [4 points]

- Put your loss & perplexities from training here, both before and after hyperparameter tuning.

Before hyperparameter tuning: Training Loss: 5.1886. Validation Loss: 5.1310.
Training Perplexity: 179.2245. Validation Perplexity: 169.1792.
After hyperparameter tuning: Training Loss: 4.4945. Validation Loss: 4.4542.
Training Perplexity: 89.5250. Validation Perplexity: 85.9847.

- Explain what you did here as well, you can use another slide if needed
- I use LSTM, which performs better than RNN. Then I add a weight decay (1e-5) parameter to optimizer to prevent overfit. Also I set epoch to 30.

Transformer Results [5 points]

- Put your loss & perplexities from training here, both before and after hyperparameter tuning.
- Before: Training Loss: 4.9300. Validation Loss: 4.8563. Training Perplexity: 138.3757. Validation Perplexity: 128.5462.
- After: Training Loss: 2.9833. Validation Loss: 3.3124. Training Perplexity: 19.7528. Validation Perplexity: 27.4496.

- Explain what you did here as well, you can use another slide if needed
- I change the learning rate to 1e-2.