Assignment 3

Your name: Adrian Kukla
Your GTID: 904056948

Transformer Results [5 points]

- Put your evaluation BLEU and Exact Match Accuracy here, both before and after hyperparameter tuning.
- Before changes:
  - Exact Match Accuracy: 18/100 = 18% correct
  - BLEU = 86.72
- After changes:
  - Exact Match Accuracy: 91/100 = 91% correct
  - BLEU = 100
- Explain what you did here as well, you can use another slide if needed
  - Lowered learning rate from 5e-4 to 3e-4 for better convergence and training stability.
  - Increased batch size from 16 to 32 to allow for more stable gradients.
  - Increased number of iterations from 10k to 12k to give the model more time to learn and converge.

Transformer Results [5 points]

- Explain what you did here as well (continued…)
  - Increased model capacity:
    - Increased from 4 to 8 layers to capture more complex patterns
    - Increased from 64 to 128 embedding dimensions to be able to store richer features
    - Increased from 2 to 8 attention heads to better understand context of a word in a sentence

  - In the tokenizer, the only change made was to add "ay" and "way" as special tokens as they're common Pig Latin suffixes, which helps the model understand common linguistic patterns in Pig Latin transformation. Following doing research on Pig Latin I realised that words either have consonants moved to the end concacenated with "ay", or have "way" added to vowel starting words.