HW 10: ECE 601 Machine Learning for Engineers

Important Notes:

1. When a HW question asks for writing a code, you would need to include the entire code as well as the output of the program as well as any other analysis requested in the question.

2. Submit your solution in **one** file called `homework10_UMassUSERNAME.zip`.

Implementing a Mini Transformer Language Model

In this assignment, you will complete the implementation of a small-scale Transformer-based language model from scratch. This exercise is designed to help you understand the inner workings of attention mechanisms, positional encoding, and transformer blocks in autoregressive language modeling tasks.

You will work through a Jupyter Notebook where key components of the model are already scaffolded. Your objective is to implement the missing parts and train the model on a tiny dataset of English sentences.

- Unzip `HW_10 Due Date April 24.zip`
- Copy `ECE601-Module10` into your Google Drive
- Use the provided starter notebook: `transformer_language_model_exercise.ipynb`.
- Do not use PyTorch's built-in transformer layers; instead, implement from first principles using only `nn.Linear`, `nn.LayerNorm`, and related components.
- Include inline comments in your code to explain what each block is doing.