DS5007 Deep Learning Lab 5
TRANSFORMERS

Max Marks: 10
Deadline: 21/03/2025, 12:00 PM

Instructions

- Provide well-commented, indented code with meaningful variable names.
- Write the task description in separate text blocks before the corresponding code block.
- Carefully follow the task requirements and use only the specified libraries or approaches.
- Ensure all plots have appropriate axis labels, titles, and legends.
- Submit a single Jupyter Notebook (.ipynb) file named YourName_YourRollNo_Assignment4.ipynb.

Question 1 (7 Marks)

Fine-tune a Transformer Model for Two Subtasks: Text Summarization and Named Entity Recognition (NER)

Subtask 1: Text Summarization (4 Marks)

- Use a **pretrained model** like facebook/bart-large-cnn or t5-small.
- Fine-tune the model on a **custom dataset** (e.g., scientific abstracts, news articles).
- Evaluate the model using the **ROUGE score**.
- Generate summaries for **5 test samples**.

Subtask 2: Named Entity Recognition (NER) (3 Marks)

- Use a **pretrained NER model**, such as bert-base-cased-finetuned-conll03-english.
- Fine-tune the model on a **custom dataset** (e.g., CoNLL-2003 or custom annotated text).
- Extract **entities like PERSON, LOCATION, and ORGANIZATION** from the text.
- Report **F1 score** and **accuracy** on the test set.

Question 2 (3 Marks)

Implement Self-Attention Mechanism from Scratch

- Write a **custom PyTorch or TensorFlow implementation** of the **Scaled Dot-Product Attention**.
- Input: A **random 5x5 matrix** representing token embeddings.
- Steps:
    - Compute **Query (Q), Key (K), and Value (V) matrices**.
    - Calculate **attention scores**.
    - Visualize the attention scores as a **heatmap (optional bonus)**.

Submission Guidelines

- Ensure your code is well-structured, readable, and includes comments.
- Submit a Jupyter Notebook (.ipynb) file with all results and outputs included.
- Ensure all plots have appropriate labels and titles.
- Late submissions will incur penalties as per course policy.

References

- Use a **pretrained model** like facebook/bart-large-cnn or t5-small or any of your choice.
- Fine-tune the model on the CNN/DailyMail dataset.
- Reference Guide: Hugging Face Transformers - Summarization Task
- Dataset CoNLL-2003 dataset for subtask2 or use any of your choice
- Self-attention from scratch