Homework 6

Start Assignment

Due Tuesday by 11:59pm **Points** 100 **Submitting** a file upload
Available Apr 24 at 12am - May 3 at 11:59pm

Homework 6 (100 Points)

Homework instructions:

- You will submit: (1) a report (pdf file) that explains your answers for each problem. The report MUST contain the link to your Github repository to access the source code you have developed for this homework. Your report should contain your name, student ID, and homework number. (2) a second pdf file that contains both source codes, results, and plots you have developed and plotted for this homework. You can easily generate this pdf in Jupyter Lab or Jupyter Notebook (sample pdf provided on Canvas.). The pdf file name should contain your name, student ID, and homework number.
- In your report, provide separate and clear responses for each problem. Make reasonable assumptions where necessary and clearly state them! Be sure to show all the work involved in deriving your answers! If you just give a final answer without explanation, you may not receive credit for that question.
- You may discuss concepts with your classmates. This fosters group learning and improves the class' progress. However, make sure to submit your own independent and individual solutions.
- **Make sure to use Google Collab (with GPU addon option), or your personal GPU card for the training.**

Problem 1 (50pts)

For the problem of Machine Translation problem: Train a deeper Transformer than what we did during the lectures. How does it affect the training speed, model complexity, and translation performance both quantitatively and qualitatively? Report and plot your results.

Problem 2 (50pts)

For the problem of the Vision Transformer, we need, in lectures, to train a deeper Transformer with more multiheaded self-attention blocks. How does it affect the training speed, model complexity, and validation accuracy? Report and plot your results.