

CSE 4392 Special Topic: Natural Language Processing

Homework 10 - Spring 2025

Due: April 17th

Question 1 - 50%

Implement and Test Multihead Attention from scratch with Python. You should test it by producing equivalent results to your own computation on paper or to PyTorch's built-in Multihead Attention module (feel free to set weights to identity matrices for testing only). In a separate paper or PDF, explain what each step of computation does in this multihead attention block.

Preferably, submit a PDF notebook using this tool. Or submit a Python file and a report documenting your work and containing test outputs.

Question 2 - 50%

Revisit your code from last tutorial for fake news detection and add transformer encoder (using PyTorch) and compare it with LSTM, GRU and RNN in terms of mean inference speed over the test set and classification performance. Beware that the transformer may require more tuning and that the transformer encoder from PyTorch may not have a built-in positional encoding (it's in a different module).

Preferably, submit a PDF notebook using this tool. Or submit a Python file and a report documenting your work.