

ASSIGNMENT 2

Deadline: 11:59 pm, March 19, 2025

Submit via Blackboard with a VeriGuide receipt.

Please follow the course policy and the school's academic honesty policy.

Submission Instructions:

Students must submit a zip file containing:

- A PDF file with their written solutions.
- The implemented `Assignment2_Q2.py` file.

For Problem 2, students must also either paste the output or include a screenshot of the output in the PDF file after executing their code.

1. Compute the output of a single-head self-attention mechanism without scaling for the given input embeddings and weight matrices. The input sequence with three 2-dimensional vectors is

$$X = \begin{bmatrix} 0.8 & -0.4 \\ 1.2 & 0.3 \\ -0.5 & 1.1 \end{bmatrix}$$

and learnable weight matrices are

$$W_Q = \begin{bmatrix} 0.6 & -0.2 \\ 0.1 & 0.5 \end{bmatrix}, \quad W_K = \begin{bmatrix} -0.3 & 0.7 \\ 0.4 & 0.9 \end{bmatrix}, \quad W_V = \begin{bmatrix} 1.0 & -0.5 \\ 0.2 & 0.8 \end{bmatrix}.$$

(1) Compute the Query (Q), Key (K), and Value (V) matrices for the input sequence.

(2) Calculate the attention scores using dot-product attention and softmax function. Show all intermediate steps.

Note: Do not scale the scores (i.e., skip dividing $\sqrt{d_k}$).

2. You will implement a single-head self-attention mechanism using PyTorch. This will help you understand the fundamental principle of the self-attention mechanism.

(1) Complete the self-attention function in the attached Python file. Complete the following steps:

- Calculate Q, K, V matrices.
- Compute attention scores (dot-product).
- Apply softmax.

Note: Use `torch.matmul` for matrix multiplication and `F.softmax`.

(2) Compute scores using the `SelfAttention` class implemented in part (1):

- Set random seed to **42** for reproducibility.
- Initialize a `SelfAttention` module with `d_model=4`.
- Create an input tensor `x` with shape (2, 3, 4) (`batch_size=2`, `seq_len=3`, `d_model=4`).
- Perform a forward pass and print the scores tensor (shape and values)

Note:

- Do not modify the random seed.
- Do not import additional libraries.
- Follow the TODOs in the comments of the attached Python file for implementation guidance.

*** END ***