

Marked exercises, lecture 5

This is the fourth set of marked exercises, which are a part of Submission 1. In this assignment, you will get familiar with the attention mechanism of transformers. Detailed instructions and explanations are provided in the accompanying Jupyter notebook. In total you can get 20 points out of 60 points for Submission 1 by completing all marked exercises related to lecture 5.

1.5.1 Implementation of Scaled dot-product attention

1.5.2 Implementation of MultiHeadAttention