Heinrich-Heine-Universität Düsseldorf
Computational Cell Biology
Alexander Kroll

Applications of Transformer Networks in Bio- and Cheminformatics

Worksheet 3

Submission Deadline:	06. Mai 2024, 11:59 pm
Discussion of solutions:	07. Mai 2024, 4:30 - 6:00 pm

Submission Instructions:

- Upload a single Jupyter notebook with your solutions for Exercise 3.1 & 3.2.
- Submit your solutions by uploading the Jupyter notebook to https://uni-duesseldorf.sciebo.de/s/hCt1rTP23EeWmUC. The uploaded file should have the following filename "lastname_studentID_worksheet3".

Excercise 3.1 *Protein Transformer Network Encoder* (100 Points)

In this task, you will implement a protein transformer network encoder and you will train it for the task of predicting masked tokens in the input sequences. The Jupyter notebook *data/worksheet3/Worksheet3.ipynb* contains already many building blocks for this process. You need to follow the exercises in this notebook to complete the missing parts.

Excercise 3.2 *Use of LLMs* (0 Points)

State for which exercise you have used LLMs (large language models) such as ChatGPT or GitHub Copilot. State which tools you have used and for which steps. This answer does not influence how many points you receive for your submission.