

Protein Prediction II

Winter 18/19

for Computer Scientists
for Bioinformaticians

Timeline

25.10.	Introduction
1.11.	No exercise
8.11.	Build a machine learning pipeline
15.11	Optimize Word2Vec parameters
22.11.	Discuss W2V optimizations
29.11.	Present 'final' W2V optimizations
6.12.	No exercise, Dies academicus
13.12.	Present current progress
20.12.	No exercise session

10.1.	Feedback about your progress (no presentation necessary)
17.1.	Q&A session
24.1.	Final presentations w/ Prof. Rost Deadline for method submission
31.1.	Q&A Session Method evaluation
7.2.	Exam

Analysis of extreme predictions

- Look only at the most confident predictions
 - The top 1% and/or 10% of all residues with the highest prediction scores
 - The highest-scoring residue of each protein
- How many of those predictions are correct?
- Idea: Can we at least reliably find the approx. location of a binding site?