A U S H A N G

FREIE UNIVERSITÄT BERLIN

Fachbereich Mathematik und Informatik

Promotionsbüro, Arnimallee 14, 14195 Berlin

D I S P U T A T I O N

Mittwoch, 13. Februar 2019, 8:00

Ort: Seminarraum 005

(Fachbereich Mathematik und Informatik, Takusstr.9, 14195 Berlin)

Disputation über die Doktorarbeit von

Herrn Mathias Kuhring

Thema der Dissertation:

Computational Methods for Omics Sequence Data with Focus on non-model Organisms

Thema der Disputation:

The impact of deep learning on de novo peptide sequencing

Die Arbeit wurde unter der Betreuung von **PD Dr. B. Renard** durchgeführt.

Abstract: Deep learning is an advanced, powerful and popular area of machine learning primarily based on multi-layer artificial neural networks, which in turn were inspired by biological neural networks. It features a broad range of applications in industry and science, including but not limited to speech and audio recognition, natural language processing, computer vision, economics and finance as well as bioinformatics. Computational analysis of biological sequences utilizes deep learning more and more to advance research in, for instance, protein binding site prediction, subcellular localization, ab initio protein structure prediction and more recently de novo peptide sequencing from tandem mass spectrometry (Jurtz et al. 2017, Tran et al. 2018) With increasing data resolution and quality as well as computational performance, de novo peptide sequencing gained traction as complementary and alternative method compared to spectral library or protein database searches. As such, it facilitates the analysis of non-model organisms and environmental communities, among others, which often suffer from incomplete or unbalanced proteome databases. At the same time, de novo peptide sequencing is more sensitive to spectra quality and more affected by, for instance, low peak intensities, missing peaks and noise. (Muth and Renard 2017) By introducing deep learning to de novo peptide sequencing, DeepNovo is able to improve length and accuracy of peptide sequences as well as resulting assembled protein sequences, thereby surpassing major established methods. DeepNovo's modular implementation features a combination of state-of-the-art convolutional neural networks, recurrent neural networks as well as local dynamic programming for feature extraction and optimized sequence prediction from tandem mass spectra. With a broad range of datasets from many different species as well as technologies, DeepNovo demonstrates its advantages including exceptional robustness and ability to adapt to new data, thereby establishing its value for future applications. (Tran et al. 2017)

Die Disputation besteht aus dem o. g. Vortrag, danach der Vorstellung der Dissertation einschließlich jeweils anschließenden Aussprachen.

Interessierte werden hiermit herzlich eingeladen

Der Vorsitzende der Promotionskommission
PD Dr. B. Renard